federal register

Friday
January 25, 1980

Part II

Department of Health, Education, and Welfare

Food and Drug Administration

FD&C Red No. 2; Denial of Petition For Permanent Listing; Final Decision

DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Food and Drug Administration

[Docket No. 76C-0033]

FD&C Red No. 2; Denial of Petition for Permanent Listing; Final Decision

AGENCY: Food and Drug Administration.

ACTION: Final Decision Following a Formal Evidentiary Public Hearing.

SUMMARY: The Commissioner of food and Drugs is issuing the final decision following a formal evidentiary public hearing to consider objections to the agency's denial of the petition for permanent listing of Red No. 2 as a color additive. The Commissioner affirms the Initial Decision, which denied the petition, with the supplementation and modifications in the Final Decision.

EFFECTIVE DATE: April 24, 1980.

ADDRESS: The transcript of the hearing, evidence submitted, and all other documents cited in this decision may be seen in the office of the Hearing Clerk (HFA-305), Rm. 4-65, 5600 Fishers Lane, Rockville, MD 20857, from 9 a.m. to 4 p.m., Monday through Friday.

FOR FURTHER INFORMATION CONTACT: Richard T. Hunt, Regulatory Policy Staff (HFC-10), Food and Drug Administration, Department of Health, Education, and Welfare, 5600 Fishers Lane, Rockville, MD 20857, 301-443-3480.

SUPPLEMENTARY INFORMATION:

I. History

In the **Federal Register** of April 9, 1976 (41 FR 15053) the then Commissioner of Food and Drugs, Alexander M. Schmidt, denied a petition to list FD&C Red No. 2 (Red No. 2) permanently as a color additive. The Certified Color Manufacturers' Association (CCMA) and the Cosmetic, Toiletry, and Fragrance Association (CTFA) objected to the denial, and CCMA requested a formal hearing on its objections. CCMA is referred to as the petitioner in this Decision because it was one of the associations that filed the original petition to list Red No. 2 permanently as a color additive. In the **Federal Register** of July 20, 1976 (41 FR 29896), the then Commissioner granted a formal hearing on the objections. At the conclusion of the formal hearing, the Administrative Law Judge (ALJ) issued an Initial Decision (ID) on March 30, 1978, denying the petition for permanent listing.[1]

CCMA filed exceptions to the Initial Decision under 21 CFR 12.125[2] and the Bureau of Foods (Bureau) filed a reply to the exceptions (Reply) requesting an adoption of the Initial Decision and denial of the petition for permanent listing for all uses. CCMA requested oral argument. Ex., p. 1. Because I do not find oral argument necessary, I am denying that request. 21 CFR 12.125(e).

I am issuing this Final Decision under section 706 of the Federal Food, Drug, and Cosmetic Act (21 U.S.C. 376) and 21 CFR 12.130 to take final agency action with respect to the petition in light of the Initial Decision, the exceptions to it, and the other material in the record of the hearing. In reaching this decision, I have all the powers I would have in making the Initial Decision. 21 CFR 12.130(a). The Initial Decision and The Final Decision must be based upon a "fair evaluation of the entire record," under section 706 of the act, and should also satisfy the provisions of 21 CFR 12.120 and 12.130. After reviewing the entire record, I deny the petition for permanent listing of FD&C Red No. 2 on the basis of the Initial Decision on supplemented and modified by this Decision.

Although Red No. 2 has never been approved for permanent use as a color additive, at one time the color was permitted to be used in foods, drugs and cosmetics as a provisionally listed color additive under the transitional provisions of the color additive amendments. The agency terminated the provisional listing of the color in the **Federal Register** of February 10, 1976 (41 FR 5823), as implemented on February 12, 1976 (see the **Federal Register** of February 13, 1976 (41 FR 6774)). The termination of the provisional listing was upheld *in Certified Color Mfg. Ass'n* v. *Mathews,* 543 F.2d 284 (D.C. Cir. 1976). The correctness of the termination of the provisional listing is not an issue in this proceeding; the hearing and this proceeding concern only the permanent listing of Red No. 2 as a color additive.

This Decision will examine the three issues on which the notice of hearing granted a hearing, and the particular disputes concerning these issues as they developed at the hearing. The three issues on which a hearing was granted are:

(1) *Carcinogenicity.* The notice of hearing granted a hearing on whether data submitted by petitioners, in addition to other data before FDA, establish to a reasonable certainty that FD&C Red No. 2 is not a carcinogen in man or animals.

(2) *Chemistry.* The notice of hearing granted a hearing on whether on the facts as to Red No. 2 the considerations set forth in 706(b)(5)(A) of the act (21 U.S.C. 376(b)(5)(A)) support a determination that Red No. 2 is safe for its intended uses. At the hearing this issue concerned whether Red No. 2 should not be considered safe due to questions concerning its stability, metabolism, and purity all of which, for convenience, are described in this decision as "the chemistry issues."

(3) *Topical Uses.* The notice of hearing granted a hearing on whether FD&C Red No. 2, if not approvable for all petitioned uses, may nevertheless be approvable for certain limited uses.

Before discussing these issues, I will examine the petitioner's obligation under the statute to show the safety of a color additive, and the general exceptions to the Initial Decision raised by CCMA.

II. General Issues

A. *Petitioner's Burden to Show Safety.* Under the law, a color additive can be included in the permanent list of approved color additives only if "the data" submitted to the agency in a color additive petition "establish" that use "will be safe" (21 U.S.C. 376(b)(4)). As interpreted by FDA, an additive is "safe" under the statute only if there is "convincing evidence that establishes with reasonable certainty that no harm will result from the intended use of the color additive." 21 CFR 70.3(i).

Thus, it is clear that the petitioner seeking approval for permanet listing of a color additive has the burden of proof to demonstrate by adequate tests that the color is safe. The additive cannot be approved if there is no information about the safety of the additive, or if there is inadequate information to determine safety for use to a reasonable certainty. Thus, denial is required when the safety of the additive is unknown or uncertain, even though the available information does not affirmatively establish that there is a particular hazard posed by use.

Congress could have established a different system. Indeed, before the passage of the 1960 Color Additive Amendments, non-coal tar colors could be marketed without an advance demonstration of safety, with FDA able

[1] Testimony was submitted in written form, with an opportunity for oral cross-examination. Written testimony was given exhibit numbers. Citations to the record in this Decision are as follows: CCMA exhibits (P-); Bureau's exhibits (G-); transcript of cross-examination (Tr.,); appendix to the Initial Decision (ID App.,).

[2] The hearing was held under the procedural rules 21 CFR 2.48 et seq. (1976)) in effect at the time the notice of the hearing was issued. However, the ALJ ruled that the Initial and Final Decisions and exceptions would be governed by FDA's revised procedural rules as found in 21 CFR 12.120 et seq. (1978). Tr., pp. 3-5. I have observed the revised rules in issuing this Final Decision.

to stop sale of these colors only if FDA could show, through its own testing, consumer injuries or other means, that the color posed a hazard to health. Under such a system, until FDA could demonstrate a potential hazard an additive could be sold for use by the general public even though no safety testing had been done on the additive, or the safety testing was inadequate. With the passage of the Color Additive Amendments, Congress did not allow any color additive to be listed for permanent use until its safety had been established with reasonable certainty by the petitioner in data submmitted to the agency.

The important purpose underlying this statutory scheme was described in *Certified Color Mfg. Ass'n.* v. *Mathews, supra:*

> The Color Additive Amendments of 1960 reflect a Congressional and administrative response to the need in contemporary society for a scientifically and administratively sound basis for determining the safety of artificial color additives, widely used for coloring food, drugs, and cosmetics. The Amendments reflect a general unwillingness to allow widespread use of such products in the absence of scientific information on the effect of these products on the human body. The previously used system had some glaring deficiencies, and the 1960 Amendments were designed to overcome them. This was accomplished by the establishment of a dual system of registration: a permanent listing and a provisional listing. A color additive would be permanently listed if those desirous of producing it had proven to the satisfaction of the Commissioner that it was safe for its intended use. 543 F.2d at 286–87 (Footnotes omitted).

The petitioner's burden to show the safety of a color additive by adequate studies has a critical effect on the outcome of this proceeding. As the subsequent discussion establishes, the petitioner has not, in my judgment, shown through adequate studies that Red No. 2 is not a carcinogen. Accordingly, the petition to list the color additive for permanent use must be denied. This determination does not mean, however, that Red No. 2 has been found to be a carcinogen. The existing studies are not sufficient to show definitely either that it is a carcinogen or that it is not. The questions about its possible carcinogenicity are simply unresolvable on this record. In this type of situation, the petition must fail because it is insufficient at present to meet the statutory requirement that safety be affirmatively shown with reasonable certainty.

For several years, the agency has maintained that the existing studies are inadequate to establish the safety of Red No. 2. During this period, the petitioner could have conducted further studies, which might have resolved the outstanding questions and might have established the safety of Red No. 2. The petitioner has not submitted any new studies in the petition, and instead has exercised the statutory right to obtain a full evidentiary hearing to test the correctness of FDA's original determination that the existing studies are inadequate. This decision completes the proceeding to review the adequacy of the existing studies. After reviewing the hearing record, I agree with the previous determination of the agency that the existing studies are inadequate to establish the non-carcinogenicity of Red No. 2.

If, at some future point, the petitioner conducts a further study or studies that adequately demonstrate the non-carcinogenicity of Red No. 2, the petitioner can submit a new petition to seek approval of Red No. 2. The agency will examine any new study or studies carefully, and will approve the petition if a sufficient showing is made that Red No. 2 is not a carcinogen and is otherwise safe.

In reviewing the adequacy of the existing studies, I have, in accordance with the philosophy of the color additive law, adopted a conservative approach in order to be sure that the public health will be adequately protected. Thus, in considering the level of statistical significance or the statistical procedure (such as a one-tail or two-tail test) to be used to evaluate the studies on Red No. 2, I have used methods that are valid and are also the ones most likely to detect any carcinogenic effect that may be present. In a regulatory proceeding such as this, the choice among statistical procedures is not an abstract or academic matter, but rather should be guided by the purpose for which a study is being used. When a study is used to evaluate the safety of a substance to be widely used by the public, the risk of a false negative—of incorrectly failing to detect an adverse effect that is present—is of greater concern than the risk of a false positive—of incorrectly reporting an adverse effect when none exists. Although it is not necessary at this point to state a position on the procedures to be used in every case, it is clearly appropriate in the case of Red No. 2 to adopt a conservative approach. The studies on the safety of Red No. 2 have deficiencies that weaken the confidence to be place in their results, and that make it more difficult to detect positive effects. Some statistical analyses of the tumor findings in one of the studies suggest possible cancer findings. In this type of situation, it is appropriate to rely on statistical analyses that use methods and procedures that will increase the likelihood of detecting any carcinogenic effect the color may induce. Moreover, the present record discloses additional reasons for being concerned about the possible carcinogenicity of Red No. 2, including the results of the Andrianova (Russian) study and the short-term screen studies. These questions reinforce the concern about safety arising from the other factors.

I am not, however, imposing an absolute standard of safety for the evaluation of safety studies. As discussed in the following section of this decision, the statute establishes a requirement that safety be shown with reasonable certainty. I would not use a procedure, even if it were the most conservative, if the procedure were not a valid one. If the questions about a substance or the defects in a study are insubstantial, they do not preclude approval of the substance. However, when uncertainty remains about safety, after a fair evaluation of the record in accordance with scientific principles of evaluation, then, under the applicable law, the importance of protecting the public health must guide the final decision. It is appropriate to be cautious and prudent in determining whether safety has been adequately shown. The use of a conservative approach in evaluating studies and in selecting the statistical methods to be used ultimately reflects the fact that Congress has imposed upon the petitioner the burden to establish safety with reasonable certainty. The record in this proceeding does not permit me to conclude that the safety of Red No. 2 has been established.

B. *General Exceptions.* The first 24 pages of the 141 pages of exceptions to the Initial Decision filed by CCMA concern what CCMA describes as "general areas of exceptions." CCMA maintains that the Initial Decision failed to provide a "fair evaluation of the record," in that the ALJ unfairly evaluated the evidence and distorts petitioner's testimony. CCMA also excepts, at length, to the ALJ's failure to apply the proper legal standards to the evidence and issues" in the proceeding. Ex., pp. 25–57, at 25.

The petitioner asserts that the ALJ in practice applied an absolute standard of proof of noncarcinogenicity, under which FD&C Red No. 2 could not be found safe unless every possibility regarding cancer was entirely disproved. Ex., pp. 18–22, 36–40. In addition, CCMA argues that the ALJ "all but ignored" certain memoranda because they were

hearsay and the authors involved were not available for examination. Ex., pp. 40–43. Finally, CCMA excepts to the ALJ's failure to strike from the hearing matters relating to the chemistry of Red No. 2 that were not expressly listed in the notice of hearing. Ex., pp. 43–58.

C. *Analysis.* I have decided, for the reasons discussed in section IV of this Decision, that it is not necessary to resolve the chemistry issues in this proceeding. With respect to the hearsay issue, I believe that it is inappropriate to exclude information or give it virtually no weight solely on the basis that it is hearsay. Under the rules in effect at the time of the hearing, hearsay evidence is not automatically excluded, but any information may be excluded if it is not reliable. 21 CFR 2.81(c) (1976). The ALJ decided that "as hearsay" certain evidence was unreliable and therefore should be excluded. ID, p. 9. In addition, the ALJ gave little weight to memoranda stating the opinions of some Bureau scientists concerning one of the studies of Red No. 2 apparently because the memoranda "suffer from the unreliability inherent in hearsay evidence." ID App., p. 34.

Although unreliable evidence should not be given weight, not all hearsay evidence should be deemed unreliable, as the ALJ seems to have believed. Instead, the information should be evaluated on its merits and given appropriate weight. The absence of the writer to provide further information on the basis for the statements, and the effect of changing circumstances, should be taken into account as affecting the weight to be given to the evidence. I have followed this practice in evaluating the record in this proceeding.

CCMA's other general objections essentially parallel the specific exceptions made by CCMA and recapitulate its major points. To evaluate whether there is any merit to the contentions that important evidence was ignored, or other evidence distorted, it is essential to understand each specific issue raised and all the relevant information on it. Consequently, I have considered these general exceptions in evaluating the specific exceptions, and I have discussed the general exceptions as pertinent to the specific issues.

An illustration of the need to examine the specific issues is provided by CCMA's general objections to the failure of the ALJ to evaluate the record properly concerning the significance of the negative results of the Ames test, a short-term screening study for carcinogenicity. CCMA objected to the ALJ's statement that there is a "general uncertainty of the conclusiveness" of the Ames test in light of the estimates of 90–95 percent reliability of this test made by three of CCMA's witnesses and one of the Bureau's witnesses. Ex., pp. 6–8 quoting ID App., p. 2. It is clear, however, that the Initial Decision considered the estimates of reliability made by these witnesses, as well as the 65 percent estimate of reliability made by another of the Bureau witnesses, and the reservations of one of petitioner's witnesses about short-term screening studies (P–150(2), pp. 42–43) since these estimates and related evidence are specifically discussed in the paragraphs preceding the reference to "general uncertainty of the conclusiveness of the test." ID App., p. 2. This statement seems to have been a summary way of indicating, without specifying the exact degree of uncertainty, that there was enough uncertainty that it would be unjustified to regard the negative results as showing that Red No. 2 lacks carcinogenic potential. Thus, the Initial Decision evaluated the evidence, but weighed it differently than CCMA would like. Moreover, the Initial Decision discusses a "complicating factor," which makes it less likely that the Ames test would give reliable results for a substance like Red No. 2 that is metabolized in the gut. ID App., p. 2. This complicating factor is not discussed in CCMA's general exceptions (Ex., pp. 6–8), but is addressed, as it is in this discussion, in connection with the specific issues on short-term screening studies and their bearing on carcinogenicity. Ex., p. 125.

Similarly, the CCMA objection that an absolute standard of safety was required has to be evaluated in conjunction with the specific issues. The ALJ stated that he was observing the standard all regard as the correct one, i.e., whether Red No. 2 has been shown to a reasonable certainty not to be a carcinogen. ID, pp. 7–8. Indeed, the ALJ stated that there is "no requirement of absolute certainty * * * because science is unable to assure complete safety." ID, p. 8. CCMA argues, nonetheless, that, although the ALJ stated the correct standard, in fact he applied an absolute standard. Ex., pp. 18–22. The nature of reasonable certainty cannot be specified with exactitude, however. Whether absolute certitude was in fact required is a determination interrelated with the evaluation of the evidence on the specific issues concerning carcinogenicity.

CCMA argues that affirmative tests to show the noncarcinogenicity of a color are required under the law only if the general safety testing shows "some indication of inducing cancer." Ex., pp. 32–40 at 33. CCMA cites an HEW statement:

> The scientific tests that are adequate to establish the safety of an additive will give information about the tendency of an additive to produce cancer * * *. Any indication that the additive may thus be carcinogenic would * * * restrain [FDA] from approving * * * the additive unless and until further testing shows to the point of reasonable certainty that the additive would not produce cancer * * *. Ex., p. 33 citing 104 Cong. Rec. 17415.

Since, in CCMA's view, the 1959 Webb rat feeding study is adequate to show the safety of the color, CCMA believes the ALJ improperly imposed a burden to disprove every possibility of cancer. Ex., pp. 35–36. However, the cited statement by HEW points out the importance of having adequate general safety testing, both to show general safety and to evaluate the carcinogenic potential of an additive. CCMA's argument does not take into account either the inadequacies of the Webb study, as judged by current standards, to evaluate the safety and carcinogenic potential of the additive, or the significant questions about the carcinogenicity of Red No. 2 that have arisen from subsequent information, including the Taylor/Monlux study, the Andrianova study, and short-term screening studies. The evaluation of these studies is considered in detail in the next section of this Decision, which relates to the carcinogenicity issue.

Finally, it should be noted that it is not necessary to discuss every item of evidence in order to have evaluated it adequately. The decision can properly focus on the evidence important to the outcome. It is also clear from the detailed summary in the Appendix to the Initial Decision that the ALJ examined the record in detail. The main text of the Initial Decision is, however, brief in its explanation of the reasons for resolution of the scientific issues. In reaching this Final Decision, I have evaluated the entire record, and discussed the principal issues and their resolution. I have concluded, after a careful review of the record, that the ultimate findings and outcome of the Initial Decision with respect to the carcinogenicity issue are essentially correct and should be upheld. Some errors were made in the Initial Decision, and its phrasing could be improved and more explanation given of its analysis. The complaints about distortion of the record largely relate, however, to relatively minor differences in descriptions of the parties' positions, and reflect CCMA's disagreement with the weight given by the ALJ to the

evidence. When the record is evaluated on an overall basis, the limitations of the Initial Decision do not undercut the validity of its basic finding that Red No. 2 has not been shown with reasonable certainty not to be a carcinogen. Consequently, I affirm the Initial Decision with the supplementation and modifications made in this Decision.

III. Carcinogenicity Issue

The ALJ found that the data presented at the hearing did not establish to a reasonable certainty that Red No. 2 is not a carcinogen in man or animals. ID, p. 20. In his general discussion of the safety testing done on Red No. 2, the ALJ stated that there were "many negative results and several equivocally positive ones." ID, p. 19. Despite the long use of the color and an apparent reasonableness "[o]n the surface" in considering approval, there were several "complicating factors." These include the fact that the color belongs to a class of chemicals many of which are carcinogenic, and that carcinogens, as well as mutagens, have a long latency period and may be missed despite long use. ID, p. 20. He concluded:

> There are numerous unanswered questions concerning the safety of Red No. 2 raised by the test data presented on the record. The total picture presented is one consisting of many positive findings which have not been satisfactorily explained. Under such circumstances, the evidence of record in this proceeding cannot be considered as constituting the requisite showing necessary to an ultimate finding of safety for Red No. 2. ID, p. 20.

Although the ALJ relied on the "total picture," rather than a single study, for the conclusion that the safety of Red No. 2 had not been adequately established, for purposes of evaluating the record and the exceptions it is useful to discuss separately each of the important studies and items of information bearing on carcinogenicity.

A. *Taylor/Monlux Study (p–6,p–7).* 1. *Introduction.* The 1975 FDA chronic feeding study (P–7), called the Taylor/Monlux study in the Initial Decision and in this Decision, was a well-designed feeding study on rats conducted under the direction of Dr. Jean M. Taylor, an FDA toxicologist. The histopathological evaluation was made by Dr. William S. Monlux and other scientists at FDA. P–6. This study involved feeding a total of 500 rats at a control level and 4 dosage levels (.003, .03, .3 and 3 percent). The scientist who conducted the study found that this study revealed "minimal effects" of the test compound and no adverse effect on weight gain or survival. P–7, p. 27. The report also states:

> The histopathological lesions, both neoplastic and non-neoplastic, were characteristic of the aged rat and seen in both control and test animals. Incidence and severity of the lesions could not be related to administration of the test compound. ID.

However, there were defects in the execution of this study, consisting of a mix-up of animals during the study, and a "high" loss of animals to autolysis. P–7, pp. 23, 29. (Autolysis is a rotting of tissues that begins shortly after death and that makes examination of tissues more difficult.) The FDA Toxicological Advisory Committee (TAC) formally agreed that the Taylor/Monlux study was of such quality that it could not be used to demonstrate the safety of Red No. 2. Minutes of the TAC meeting, March 8–9, 1976, are on file with the Hearing Clerk, and of which I take official notice. In addition, a statistical analysis of the tumor findings in the study, made after its completion by Dr. David W. Gaylor, a member of the TAC and the chief of biometrics at the National Center for Toxicological Research, indicated an increase in a variety of malignant neoplasms among aged female rats. G–226 p. 21. These findings concerning the Taylor/Monlux study were a factor in the agency's denial of the petition for permanent listing of Red No. 2 in the **Federal Register** of April 9, 1976 (41 FR 15053–15054). FDA stated that:

> [T]he Adverse (sic) implications of the [Taylor/Monlux] study cannot be ignored. In light of continuing public concern and the serious new questions of carcinogenesis raised by the [Taylor/Monlux] study, the Commissioner concludes that a study or studies adequate to dispel all such questions must be performed before the color additive can be found to be safe * * *

2. *Overall Findings Concerning the Taylor/Monlux Study* —(a) *Initial Decision.* The ALJ found that this study "demonstrated a carcinogenic effect from the ingestion of Red No. 2." ID, p. 16. The specific contentions of the parties concerning the study were not discussed in the main test of the Initial Decision, but the Appendix (ID App.) to the Initial Decision described the different conclusions that the parties believed the study supported:

> This study is cited by CCMA as demonstrating that Red No. 2 is not a carcinogen (Brief at 172). However, the Bureau characterizes the results as disturbing positive evidence of carcinogenicity (Brief at 26). ID App., p. 26

(b) *Exceptions.* CCMA maintains that the ALJ "entirely ignored" the opinions of scientists, including the FDA scientists who conducted the study and others, that the results of the study are negative. Ex., pp. 64–69 at 69. According to CCMA, the study met FDA's general guidelines and other generally accepted criteria, for evaluating the adequacy of a study. Ex., pp. 69–70. CCMA also states there is "no biological support" in the record for the ALJ's conclusion that this study demonstrates a carcinognic effect, nor any explanation for the finding in the Appendix "as to what constitutes the supposed 'disturbing positive evidence of carcinogenicity.'" Ex., p. 80.

The exceptions also criticize the "lack of organization" of the Initial Decision, and the "failure to make any express choice between the positions of the parties" with respect to the adequacy of this study. Ex., p. 80.

(c) *Bureau's Reply.* The Bureau maintains that the Taylor/Monlux study cannot provide a reasonable assurance of the noncarcinogenicity of Red No. 2, despite the adequacy of the design of the study, because there are deficiencies in some aspects of the execution of the study, notably the mix-up of animals in the study and the autolysis of animals. Moreover, presumptive evidence of carcinogenicity exists due to statistical analyses of the results of the study made by Dr. Gaylor and Dr. Marvin Schneiderman. G–226, G–227. The Bureau also argues that the opinions of FDA scientists were properly given "virtually no weight" by the ALJ, because the opinions were stated before the statistical analyses became available, and because the views were hearsay and CCMA failed to call the scientists as witnesses. Reply, pp. 12–14; Brief, pp. 26–44.

(d) *Analysis.* I believe that consideration and some weight should be given to the conclusion of the scientists from FDA who conducted the study, and others who may believe that its results are negative. These views should not be rejected solely as hearsay. Neither are these views controlling, however, when questions have been raised about the adequacy of the study. I have considered the views of the scientists cited by CCMA, but I believe it is also important to consider the contentions raised by the Bureau about the defects in the execution of the study and the presumptive evidence of carcinogenicity arising from the statistical analyses of its results. These specific points are examined below, as well as CCMA's disagreement with them. This discussion of the specific disputes will provide a format for examining the views of CCMA's experts that the defects in this test are too minor to invalidate the study, and that the presumptive evidence of carcinogenicity cited by the Bureau is statistically and biologically insignificant.

Before turning to these specific points, I should note some general matters. I believe that the Initial Decision is unclear when it states that this study "demonstrated a carcinogenic effect." ID, p. 16. If the ALJ meant that the Taylor/Monlux study establishes the carcinogenicity of Red No. 2, I reject the finding. As noted above, in denying the petition to list Red No. 2—the action that led to this hearing—FDA took the position, not that Red No. 2 was proven to be a carcinogen, but that there was questions about its carcinogenicity that needed to be resolved by a further study or studies before its safety could be established with reasonably certainty. At the hearing, the Bureau maintained that the statistical analyses of this study by Drs. Gaylor and Schneiderman created presumptive evidence of carcinogenicity, which needs to be explored in further studies. The Initial Decision's finding of a demonstrated carcinogenic effect and disturbing positive evidence of carcinogenicity appears to relate to this showing by the Bureau; and, as thus understood, I accept the basic finding.

I agree with CCMA that the Initial Decision suffers from organizational defects. A specific finding with respect to each of the major contentions involved in the hearing would have provided a better focus for exceptions, and for my analysis in reaching a final decision. The general nature and organization of the exceptions have also complicated their analysis, and made it difficult to identify the specific objections being made. As a result, the final decision has had to be more detained than it would ordinarily have been.

3. *Mix-up of Animals as a Defect in the Taylor/Monlux Study*—(a) *Initial Decision.* The Appendix states that "the extent of the mix-ups is unknown" and that "it would appear that the feeding mix-up reflects adversely on the calibre of this study." ID App., pp. 27, 28. The Appendix also referred to Dr. Taylor's hypothesis that assuming "the worst possible case," that is that the mix-up involved both controls and high dose animals, lasted a month and occurred near the end of the study, "it would require a potent carcinogen to induce the some degree of response in the controls as in the high dose," but, according to Dr. Taylor, the results of the study did not support such a conclusion. P–7, p. 29. The Appendix rejected this hypothesis that the mix-up was very limited in extent because the hypothesis assumed a dose-response effect. ID, App., p. 28. The following explanation was given:

> It may be that carcinogenicity does not present itself in a dose-related manner at the levels of exposure analyzed in this study but is observed as an all-or-nothing effect. The lack of a dose-effect profile can be an indication that the test group's dose levels did not include the dose levels necessary to delineate what may be an effect which can be observed over only a small range of doses. If this is the case in the present study, the mix-up could have a very substantial influence on the ability of the experiment to reveal a dose-related effect. ID App., p. 28.

(b) *Exceptions.* CCMA excepts to the finding that the mix-up reflects adversely on the study because (1) the only Bureau witness to comment on the mix-up was a statistician, who CCMA believes is not qualified to offer biological observations; (2) the FDA scientist who conducted the study and other FDA staff believed that the mix-up was limited to two animals misplaced for 1 week (P–7, p. 24, P–49, P–144, p. 2); (3) both FDA (P–140) and the Toxicology Advisory Committee (P–33) determined that the mix-up did not taint the results; (4) CCMA's experts testified that the mix-up was not extensive because, as one of the experts observed (P–157, p. 23), the controls had a low incidence of tumors, and, if any appreciable contamination occurred and if Red No. 2 were a carcinogen, a greater tumor incidence in the controls should have been observed; (5) a mix-up would not have obscured positive results because it was not extensive and because the high dose animals were not affected; (6) the discussion of dose-response effect in the Appendix to the Initial Decision is "simply incomprehensible." Ex., pp. 10–12, 82–84.

(c) *Bureau's Reply.* The Bureau maintains that the extent of the mix-up cannot be shown with reasonable certainty. The mix-up could only have reduced the chance of detecting carcinogenicity because the high dose group would have received less of the compound, and the control might have received some. Furthermore, the technician responsible for the feeding and housing of the test animals stated that the mix-up was extensive and long-lasting. The statements of Dr. Taylor and other FDA scientists provide an insufficient basis to ignore the mix-up due to the hearsay nature of the statements and CCMA's failure to call the scientists as witnesses. The Bureau states that the extent of the mix-up may never be known. CCMA must show beyond a reasonable doubt that the mix-up was inconsequential, which, the Bureau claims, CCMA has failed to do. Reply, p. 14–15, Bureau Brief, p. 26–27.

(d) *Analysis.* The fact that there was a mix-up in the rats appears in the report of the study and in FDA memoranda, submitted by the petitioner, describing investigations of rumors about a mix-up of animals in this study. These memoranda can be relied upon notwithstanding that they are hearsay, just as the written opinions of the FDA scientists who did the study should be given appropriate weight. The testimony of the Bureau's statistical witness (Gaylor, G–226) is also appropriately considered on this point, due to the witness' familiarity with carcinogenicity testing and because the impact of a mix-up on the ability of a test to detect an effect presents a matter on which statistics is relevant.

I have reviewed the material cited in the record to determine the basis of the differing views about the extent of the mix-up and its impact on the reliability of the study. Dr. Taylor, who was responsible for the study, referred to the mix-up, and concluded, on the basis of a "thorough study of the individual weight records," that it was minor and limited to two rats in the 94th week of the study and would not affect the outcome. P–7, P–144. She also offered the analysis discussed above, that the mix-up was limited based on the "worst possible case." P–7, p. 29.

The mix-up was described in more detail in a June 26, 1975 memorandum of meeting, which was accepted as correct by the signatures of all attending. P–41. The technician for the study gave the following explanation of the mix-up:

> [A]fter the study had been ongoing for over a year, he had noticed that, based on the numbers on the cages that certain rats were receiving improper diets. At that time he had noted that the mix-up involved the controls, the low level of 0.003%, and the middle levels of 0.03% and 0.3%. It did not involve the high level of 3%. He said that to his mind the mix-up was extensive and that he had no way of calculating how long this had been going on. P–41, p. 1.

CCMA points out in its exceptions that the technician did not state that the mix-up was long-lasting, as incorrectly reported in the Initial Decision, an example in CCMA's view of "the types of distortions of the evidence characteristic of the Initial Decision." Ex., p. 10. Instead, according to the memorandum, the technician reported that he did not know any way to calculate the duration of the mix-up. This discrepancy in the description given in the Initial Decision has no significant impact on the overall assessment of the mix-up, and certainly does not warrant designation as a "distortion" of the record.

In the same memorandum, Dr. Taylor reported that, to determine the extent of the mix-up, she had examined the weight gain and food intake data and

had "tried to arrive at points of discontuniuty, that is, points at which the rats obviously could not have gained or lost the amount of weight shown on the [computer] readouts." P–41, p. 1. The pattern of mix-ups between certain cages of animals that she detected "could not be the complete case," however, according to the memory of the technician, who believed that a different pattern of displacement occurred. P–41, p. 1. He recollected that "there was displacement not only laterally but vertically * * *. In addition, he had remembered at least two rats that were misplaced in still another rack." Id. Dr. Blumenthal told the meeting that he "could not gainsay the words of a first-hand observer whose powers of observation and long experience could not be discounted," and that unless there were other means of resolving the difficulties, he would conclude there was a mix-up of "unknown duration and unknown extent." P–41, p. 2. (Quotations are from the memorandum, and are not necessarily the words of the speaker).

I believe that I, too, must give great weight to the observations and recollection of the person most directly involved. Because the pattern of mix-up suggested by Dr. Taylor's analysis does not accord with the recollection of the technician involved, I find the extent of the mix-up uncertain, and am unable to conclude with reasonable certainty that the mix-up had no significnt impact on the ability of this test to detect any carcinogenic potential that Red No. 2 may have.

Moreover, the memorandum of meeting reports that Dr. Taylor's procedure was criticized at this meeting on the basis that "the rats were so near their growth plateau that they were all very similar in body weight. Thus, although some discontinuity had been noted, nontheless, many of the rats could have been mixed and because of the similarity in weight would not have been noted." P–41, p. 2. This criticism adds to my reluctance to rely on the body weight analysis.

The petitioner maintains that the mix-up should be viewed as limited because few controls had any tumors and, if Red No. 2 were a carcinogen and accidentally fed to the controls, more tumors would be expected. P–157, p. 23 (incorrectly cited as P–156 in Ex., p. 83). This analysis seems to rely on the historical incidence of tumors in the strain to provide a baseline for estimating whether the tumor rate was low. The circumstances of a study can affect the tumor rate, however, and that is why a well-designed study includes concurrent controls and does not rely on the historical incidence of tumors. In this case, primary reliance should be placed on the incidence observed in the actual concurrent controls. Dr. Smuckler's testimony, although not specifically directed at this particular question, provides a pertinent explanation of why it is "unwise" to ignore the tumor rate actually seen in concurrent controls and to rely instead on historical experience:

> The historical control group is subject to very differnt intercurrent conditions than are the current control group. Not only might the genetic pool change in these animals over a period of years, but the environmental conditions most certainly do, the infectious diseases to which they are subject most certainly do, and not to mention the dietary regimen with which they are maintained. G–217, p. 24; see G–226, pp. 41–43.

Since the historical rate can vary from the concurrent rate in the controls for the reasons indicated by Dr. Smuckler, it is inappropriate to assume the mix-up had no impact on the controls, based on a comparison with historical experience, in the absence of an adequate study showing with reasonable certainty no carcinogenic effect from Red No. 2.

Dr. Taylor also suggested (P–7, p. 29), as described above, that a mix-up occurring near the end of the study would have had an effect on tumor findings only if the substance were a potent carcinogen, which Red No. 2 did not appear to be from the study. This analysis assumes, though, that the "worse possible case" would involve a mix-up of both control and high dose animals but only for a month near the end of the study. However, the technician first noticed the mix-up "at some time after the study had been ongoing for over a year," and he did not know how to calculate how long the mix-up had been going on. P–41, p. 1. Thus, there is no reasonable certainty about the duration and point of initiation of the mix-up, and therefore no basis for being sure that Dr. Taylor's hypothesis reflects the "worse possible case."

CCMA also maintains that the tumor incidence in the high dose animals was not depressed by the mix-up in any way because the high dose animals were not involved. Ex., p. 83. CCMA does not cite the basis in the record for its position that the high dose animals were not involved, but the basis is probably the recollection of the technician, as described in the June 26, 1975, memorandum (P–41), discussed above, that the high dose animals were not affected. Although CCMA appears to rely on this aspect of the technician's recollection, CCMA does not credit the technician's recollection that the mix-up was more extensive than the pattern identified by Dr. Taylor from the weight records. Moreover, it should be noted that the pattern of mix-up found by Dr. Taylor involved high dose animals, according to the report given by Dr. Taylor to the Toxicology Advisory Committee:

> Dr. Taylor indicated that when she had checked all records, she could only find change cards for the high-level males which suggested that perhaps they were the only group involved in the mix-up. She also felt that the mix-up probably occurred for only two or three days and involved possibly four to five animals receiving the wrong dose. Dr. Taylor said that perhaps really only the high dose and the next to high dose levels were involved. She also pointed out, however, that the animal handler's account of what happened doesn't concur with the conclusions one can reach by examining the records. Dr. Mandel stated that one cannot eliminate the possibility that the low dose animals received a higher dose. Dr. Taylor agreed that this possibility does exist. Dr. Murphy observed that according to what has been stated, the control animals were not involved in the mix-up, contrary to the recollections of the animal handler. TAC Minutes, March 8–9, 1976, p. 4.

Thus, the possibility exists that some high dose animals were involved in the mix-up. If the mix-up affected some high dose animals, perhaps the mix-up affected other high dose animals even though the technician does not recollect their involvement. Perhaps there were two different mix-ups, the one detected by Dr. Taylor and the one recollected by the technician. The very fact that there was a mix-up creates some uncertainty about its nature, extent and duration, and the reliability of the study. The general uncertainty could be dispelled only if there were clear information to determine the bounds of the mix-up. On this record, however, it is simply impossible to determine, with the confidence the law requires, what the extent of the mix-up was and whether it has an impact of consequence on the ability of this test to detect a carcinogenic effect.

The opinions of the other experts cited by CCMA (see P–150, p. 59, P–156(2), pp. 29, 30) generally relied on what "appears" from the reading of FDA memoranda describing the extent of the mix-up, and thus these opinions add little to the analysis given above of the significance of the mix-up.

CCMA also cites in support of its view statements in memoranda of FDA staff and the TAC that the mix-up was limited and did not taint the results. Ex., pp. 65–66, 83. The FDA staff

memorandum cited by CCMA (P140),[3] written on June 6, 1975, reports that "it appears" that four animals "may have received" higher doses than they should have for one week. This evaluation may have been based on the body weight analysis done by Dr. Taylor although that analysis estimated that two animals rather than four were involved. In any case, as already discussed, I believe the recollection of the technician about a different pattern of mix-up of the test animals, described in a later memorandum of June 26, 1975 (P–41), creates uncertainty about the extent of the mix-up.

The report of the TAC (P–33, p. 1) cited by the petitioner is a November 25, 1975, internal FDA information memorandum to the Commissioner, which reports that there is "reason to believe" that the intermixing is irrelevant "since the biostatisticians would have combined control and low-dose animals, not expected to react, with high-dose animals." P–33.

The full minutes of the TAC meeting indicate that the FDA staff viewed the mix-up as being of "unknown extent," and that Dr. Gaylor, a statistician and member of the TAC, initially though that the study could be of use because the controls and low doses could be combined and compared with the high dose animals, which apparently showed no carcinogenicity. P–74, pp. 17–18. The TAC wanted more information on various matters before it reached a conclusion on carcinogenicity, however, including a statistical reappraisal by Dr. Gaylor. Id. Dr. Gaylor testified that an analysis limited to malignant tumors was suggested because of the large array of tumors observed; when Dr. Gaylor did this analysis of malignant tumors, he detected a positive effect. G–226, pp. 18–19.

The TAC, at its subsequent meeting, on March 8 and 9, 1976, formally agreed that the Taylor/Monlux study was of such quality that it could not be used to determine the safety of Red No. 2. Minutes, pp. 12, 36; 41 FR 15054. Thus, I reject the petitioner's contention that the TAC believed the mix-up did not taint the evaluation of this study. To the extent the TAC held this view, it was based on an expectation of the results of a future statistical analysis; but when the Committee reexamined the matter at its March 1976 meeting, it found, as noted above, that the study was inadequate to evaluate safety due to various defects. Moreover, a number of the members found some suggestive or evidentiary value in the data, including the statistical reappraisal of this study, to show that Red No. 2 may be a carcinogen. TAC Minutes, March 8–9, 1976, pp. 18–21, 36.

Dr. Gaylor also testified at the hearing that "[t]he extent to which this mix-up occurred is unknown. It may have been very minimal. It may have been substantial." G–226, p. 18. He also stated that "the effect of the mix-up of animals could only have been to lessen our chances of being able to detect the difference between the control and treated groups." G–226, p. 19.

I, therefore, find that it is not possible to ascertain with reasonable certainty the maximum extent of the mix-up. As a result, it is not possible to view the Taylor/Monlux study as providing reasonable assurance of the noncarcinogenicity of Red No. 2.

The petitioner has criticized the discussion of the dose-response relationship in the Initial Decision. The discussion in the Initial Decision may be based on Dr. Gaylor's testimony that, while a carcinogenic effect "generally" is dose-related, for several reasons, "this is not always the case." G–226, pp. 35–37 at 36. Although the analysis in the Initial Decision about the significance of a mix-up has some merit if Red No. 2 is the type of substance that does not exhibit a dose-response effect, I am independently relying on the analysis given above in this Decision. The analysis in this Decision provides an adequate basis for believing that the mix-up was of an unknown, but possibly significant, extent, even assuming the more usual case that any carcinogenic effect from Red No. 2 would appear as a dose-related effect.

4. *Autolysis as a Defect in the Taylor/Monlux Study.*—(a) *Initial Decision.* As noted in the Appendix to the Initial Decision, the pathology report for this study stated that the majority of the interim animals "were considered to be unsuitable for detailed histopathologic examination." ID App., p. 31. (Histopathological examination is examination through a microscope.) The Appendix also states, without a citation to the record, that the pathologists "indicated that the autolysis * * * did not affect their ability to determine any carcinogenic influence of Red No. 2." Ibid. The Initial Decision also found that the effect of autolysis "appears to be similar to the effect of the feeding mix-up in that autolysis interferes with accurate diagnosis. This would serve to strengthen any positive findings * * *." ID App., p. 32.

(b) *Exceptions.* CCMA excepts to the Initial Decision for finding that autolysis interferes with accurate diagnosis and strengthens positive findings, because it constitutes "[a] rejection of the FDA pathologists' judgment as to whether they could evaluate the tumors." Ex., pp. 9, 80–82. CCMA relies on the testimony of its experts that the study was adequate, and argues that autolysis is not necessarily an all-or-nothing matter and that an adequate pathologic examination may be possible despite some autolysis. Ex., pp. 66–67, 72.

(c) *Analysis.* There is no citation to the record in the Initial Decision or the exceptions for the statement that the FDA pathologists indicated that autolysis did not affect their ability to determine a carcinogenic influence. This indication by the pathologists seems not to have been an affirmative statement by them, but an inference that CCMA maintains should be drawn because the FDA pathologists reported detailed tumor figures and did not express concern that the degree of autolysis precluded them from reaching conclusions. CCMA Brief, p. 164.

The fact that the pathologists reported autolysis indicates that they considered autolysis of some relevance in evaluating the findings. In their conclusion they also stated that "[b]ased on histopathologic findings, no apparent effect" was produced. P–6, p. xi. Thus, the pathologists' conclusion about the study seems to be limited to the tissues available for microscopic examination. The need to have microscopic examination to detect certain tumors and to be sure about diagnosis also limits the reliance to be placed on findings made without a histopathological examination.

The National Cancer Institute (NCI) guidelines call for detailed histopathological findings on all control and treated animals in chronic studies. G–11, p. 55. Thus, the guidelines emphasize the importance of avoiding autolysis in order to be able to conduct histopathological examination.

A reading of the pathologist's report leads to the conclusion that the autolysis in this study was a serious problem. Although a histopathological examination was made in this study of tissues from 450 rats, 354 of these rats were "interim" animals, and a majority of these rats simply died during the interim period without being intentionally sacrificed. As the pathology report states,

"almost all of the tissues from dead [interim] rats were found to be *in an advanced state of autolysis and were considered to be unsuitable, for detailed histopathologic examination.* From these rats, only the livers, kidneys, and gross tumor masses were

[3] CCMA cites Exhibit P–40 at page 65 of its exceptions, but this reference appears to be in error since Exhibit P–140 is cited at page 83 for the same point.

processed for histopathologic observation." P–6, p. vi. (Emphasis added.)

CCMA argues that autolysis is not an all-or-nothing matter, and that the proper procedure is to look at whether the tissues were actually readable. Ex., pp. 67, 72. Although this may be the case in some studies, the references to an "advanced state of autolysis" and a "high" loss to autolysis in the reports for this study (P–6, P–7) make it unreasonable to assumef that the autolysis in this study had no impact on the evaluation of tissues that had autolysis.

The remaining question is whether the number of animal tissues suffering from autolysis was sufficiently great to affect the adequancy of the test to detect a carcinogenic effect. The report of the study states:

> The number of rats lost to autolysis was high (see Tables 8 and 9), but as shown on Table 9, there was an average of 20.8 males and 31.4 female/level which had survived more than 78 weeks and were available for full histopathological study. P–7, p. 23.

In connection with the evaluation of another study at the hearing, there was testimony about the importance of having adequate histopathological evaluation. See pages 69–72 of this Decision. The testimony illustrates as well the impact that autolysis can have on the adequacy of a study when it prevents histopathological evaluation.

The autolysis in this study does not appear to have been biased since autolysis affected controlled and dosed animals proportionately. G–226, p. 39. The ability of the test to detect an effect is reduced, however, to the extent animal tissues cannot be fully examined to detect whether there was a greater effect in test animals. Ibid. The effect is the same as if a smaller number of animals has been used; a smaller test has less power to detect a carcinogenic response. Dr. Gaylor stated that "* * * because of the problems of autolysis, only about ⅔ of the animal tissues could be meaningfully examined. Therefore, the statistical significance level reported (in his analysis) is conservative." Ibid. Moreover, at the TAC Meeting, Dr. Monlux stated that there was "an average autolysis rate of 56 with some groups having an incidence of 100. This seriously hampered the accuracy of the diagnoses thus negating the value of any statistical approach." TAC Minutes, p. 6 (March 8–9, 1976).

I conclude that the Appendix to the Initial Decision was correct in finding that the autolysis in this test reduced its ability to detect tumors, and any positive findings could have been stronger if no autolysis had occurred, and more tissues had been available for full examination.

I recognize that the scientists who conducted the Taylor/Monlux study reported tumor findings for Red No. 2, and reported no observed differences between treated animals and controls in the incidence of lesions. P–6, P–7. However, the reports were based on the "apparent" findings in the histopathologic evalulation. The autolysis in this study, in my view, reduces the confidence to the placed in the adequacy-of the study, and particularly so in view of the suggestive positive findings found in the statistical analyses discussed below. These statistical findings could possibly have been stronger if the study's power to detect a carcinogen had not been impaired by extensive autolysis.

5. *Presumptive Evidence of Carcinogenicity Based on Statistical Analysis of Taylor/Monlux Study.*—(a) *Introduction.* Several statistical analyses were made of the tumor findings in the Taylor/Monlux study, but two are of major interest: The first is an analysis of total tumors of various types in "terminal" female rats, i.e., those that survived to the end of the study. (Total Tumor Analysis.) This analysis was initially made by Dr. Gaylor in the fall of 1975 and subsequently revised to take account of revised tumor findings based on a review of the pathological work. The initial and revised analyses of Dr. Gaylor were discussed at the Toxicological Advisory Committee meetings, described above, and were taken into account in the FDA order denying the petition for permanent listing of Red No. 2. The revised tumor findings upon which this statistical analysis was based found no malignant tumors in 14 control terminal female rats, but malignant tumors in 6 of the 21 rats in the high dose feeding level of this group. G–226, p. 21.

The other statistical analysis of major interest was made of lymphosarcomas in the femal rats. ("Lymphosarcoma Analysis".) The percentages of female rats with lymphosarcomas were 4.5, 0, 6.7, 13.6 and 11.4 in the control, .003 percent, .03 percent, 0.3 percent, and 3.0 percent feeding levels, respectively. G–226, p. 25. Both Dr. Gaylor and Dr. Schneiderman testified about the importance of this analysis at the hearing. G–226, pp. 25–26; G–227, p. 18.

(b) *Initial Decision.* The Initial Decision found that the Taylor/Monlux study "demonstrated a carcinogenic effect from the ingestion of Red No. 2." ID, p. 16. I have discussed this finding in the Initial Decision in part III A(2)(d) of this Decision. The basis for this finding is not cited, but it apparently is the statistical analyses of total tumors and lymphosarcomas in female rats made by Dr. Gaylor (G–226) and Dr. Schneiderman (G–227). These analyses showed, according to the Bureau, "[d]isturbing positive evidence" of carcinogenicity. Bureau Brief, p. 26. The Initial Decision also contained a discussion of the reasons for accepting the statistical methods used by these witnesses and rejecting the statistical methods, such as a two-tail test, that CCMA regarded as appropriate. ID, p. 17.

(c) *Exceptions.* CCMA excepts on several grounds to the finding that this study has some positive evidence of carcinogenicity:

(i) *Toxicologists' Conclusions.* CCMA maintains that the ALJ ignored the conclusions of the FDA scientists who conducted the study and the CCMA expert witnesses that the study showed the noncarcinogenicity of Red No. 2. Ex., pp. 4, 18, 40, 66–69. These experts pointed out that the tumors found are not those usually associated with carcinogens, but are invariably found in old rats. Ex. pp. 66–68. Further, the high dose treated animals lived longer than the controls; if Red No. 2 were a carcinogen, it should have shortened survival time. Ex., p. 69.

(ii) *Tumor Rate in Terminal Females.* CCMA notes that the high dose female rats had essentially the same rate of cancer as seen in the other treatment levels, and the rate appears significant in the Bureau's total tumor statistical analysis only due to the unusual occurrence that the contorls had no tumors. Ex., pp. 78–79. According to CCMA, the absence of malignant tumors in the controls indicates an inadequately reported study or an aberrant control group. Ex., pp. 72–74, 78–79.

CCMA also believes that the Initial Decision distorted CCMA's position, since CCMA does not contend, as the Decision suggests, that historical controls should be substituted for concurrent controls. Ex., p. 14–15. Instead, CCMA believes that historical controls should be examined "with care" if the historical rate differs from the concurrent controls. In effect, CCMA seeks substitution of historical controls for the concurrent controls for purposes of analysis, but is not suggesting having no concurrent controls. This meaning, however, is clear from the Initial Decision.

(iii) *Site Specificity.* CCMA also contends that the study does not have the type of evidence "usually associated" with carcinogens, since there was not a target organ effect. Ex., pp. 8–9, 73, 84–85.

(iv) *Lymphosarcoma Analysis and Different Pathological Screens.* The statistical finding made by Dr. Schneiderman and Dr. Gaylor of a linear trend of lymphosarcomas in female animals—the only site specific target organ effect claimed by the Bureau—was made on the basis of different pathological screens. Ex., pp. 16–17, 88. Moreover, Dr. Schneiderman made 16 analyses of the data, and this lymphosarcoma finding is the only statistically significant positive finding he made, and many of the other analyses are negative. Ex., pp. 77–78, 86–87.

(v) *Statistical Methods in Total Tumor Analysis.* The Total Tumor Analysis in female animals made by Dr. Gaylor used the Fishers Exact Test, a test developed for use with "fixed ratios and columns," but this study did not have fixed ratios and columns. Ex., p. 77. Although it does not state it as a specific exception, CCMA also complains that the statistical procedure used in this analysis differed from the procedures originally used after the second pathological review reported fewer tumors than found in the first review. Ex., pp. 16, 77 n. 42. In its general exceptions, CCMA also maintains that the Initial Decision incorrectly reported that CCMA regarded a .05 level of statistical significance as an absolute measure of significance. Ex., p. 17. Instead, CCMA views Dr. Gaylor's findings of statistical significance weaker than .05 as insufficient to overcome the findings of the biological scientists. CCMA also excepts to the ALJ's concern with a rising total tumor rate in this study because the various analyses made by Dr. Schneiderman did not show statistically significant results other than for lymphosarcomas. Ex., pp. 86–90.

(d) *Analysis.*—(i) *Toxicologists' Conclusions.* I agree with the Initial Decision that the opinion of the toxicologists conducting the study that the test had negative results is not dispositive. That conclusion can be rebutted by other evidence, for example, evidence showing defects in the execution of the study, such as a mix-up of test animals and autolysis, which limit the confidence to be placed in the test, or a statistical analysis showing some positive results that need further investigation.

CCMA, at several points, criticizes the Bureau's position because it is based on statistical rather than biological evidence. Ex., pp. 15–18, 59. However, the statistical analysis relied on by the Bureau is not an abstract study but an analysis of the Taylor/Monlux study and its biological results. The analysis takes the biological findings as a given, and evaluates them by a mathematical procedure to determine trends and patterns not necessarily apparent to scientists who are not statisticians. A decision about safety involves consideration of all the relevant scientific information, and a valid statistical analysis of test results has an appropriate role in safety evaluation.

CCMA makes a similar general criticism that the Bureau's evidence in this proceeding consists of statistical evidence and is not supported by biological evidence, by the testimony of toxicologists, or by the views of the FDA scientists who did the studies. As discussed above, statistical analysis of a biological study is relevant and important information. Moreover, CCMA has the burden to show with reasonable certainty that Red No. 2 is not a carcinogen. In other words, the Bureau does not have to prove that Red No. 2 is a carcinogen; it is enough for the Bureau to show that the carcinogenicity of Red No. 2 has been placed in substantial doubt and that that doubt has not been resolved. If the data are inconclusive, CCMA has failed to make the necessary showing.

Furthermore, in support of its position the Bureau has introduced pertinent testimony of "biological" experts. Dr. Emmanuel Farber, a doctor, biochemist and specialist in pathology, testified that the variety of tumors seen in this study raises a suspicion of cancer and that one could not determine from this study whether or not Red No. 2 is a carcinogen. G–219, p. 7. Dr. Edward Smuckler, a pathologist and member of the Toxicology Advisory Committee, testified that this study "did not prove either safety or lack thereof," because the "incidence of intercurrent renal disease was particularly high." G–217, pp. 23–25. Dr. Smuckler also noted that the number of surviving animals was small, but the size of the surviving group reflects the design and conduct of this study.

The ALJ made the following relevant comment about the petitioner's obligation to show the safety of a color additive:

> Throughout this proceeding, CCMA has harbored the misconception that it is incumbent upon the Bureau to provide definitive studies to establish the lack of safety of Red No. 2. In its brief, CCMA quotes Dr. Smuckler as indicating that although the available data does not indicate the safety of Red No. 2, it also doesn't indicate that it's not safe. (Brief at 170 and 215). However, as previously indicated, the statute and its history indicate that a manufacturer seeking FDA certification for a color additive must prove its safety. Old studies have proven inconclusive in this matter, and more recent studies have raised serious doubts as to the safety of Red No. 2. ID, p. 19.

It could be that there are some scientists (perhaps including even some FDA staff scientists) who might conclude from the present information that Red No. 2 is not a carcinogen. The law requires a reasonable certainty about the absence of harm, and intuitive impressions are not sufficient if not adequately supported. The ultimate decision about the safety of a color is to be made by the Commissioner in light of the governing law, the evidence of record and his expertise, and is not to be made by those in a particular scientific discipline or by certain members of the staff. After reviewing the record, I believe that Dr. Schneiderman stated well the approach to be taken in this type of circumstance:

> I think the intuitive impressions of the pathologist ought to be paid attention to, and I think the results of the statistical computations ought to be paid attention to. I think we ought to find out on what basis the pathologist's impressions are developed and on what basis the statistics are developed and see if these can't be reconciled. In the end, though I think gut conclusions should be supported by the numbers. An intuition based on objective fact should carry much more weight. G–227, p. 20.

CCMA also maintains that the statistical findings should be overlooked because the biological scientists recognized that the rats had "garden-variety" tumors characteristic of aged rats, and the treated high dose rats lived as long as the controls. Ex., p. 84. The ALJ and the Bureau correctly point out that CCMA is wrong if it maintains that a substance should not be a considered a carcinogen simply because it causes an increase in a type of cancer that occurs epidemically in older animals or humans. A CCMA witness agreed on cross-examination that it is "an important fact" if a certain agent increases the incidence of a disease commonly found in old people. Tr., p. 251–52. This citation is supportive of the Bureau's and ALJ's position, despite CCMA's argument (Ex., p. 84) that no pertinent statement was made.

(ii) *Absence of Malignant Tumors in the Control Female Terminal Animals.* CCMA maintains that the "common scientific response" when a control group has an absence of malignant tumors is to believe "either (a) the study was not adequately reported, or (b) the control group was an aberrant group * * *" Ex., p. 74. On this basis, CCMA argues that no weight should be given to the statistical finding of a higher incidence of total tumors in high dose

female terminal animals as compared with the unusual absence of tumors in the controls in the total tumor statistical analysis made by Dr. Gaylor.

I disagree. If the absence of tumors indicated that the study may have been inadequately reported, it provides an additional ground for finding that the study is not adequate to show the noncarcinogenicity of Red No. 2 and that another study is needed to resolve the question.

I also believe the record adequately supports the conclusion that the actual findings observed in the concurrent controls should be relied on for comparison purposes. G–217, p. 24; G–226, p. 41. An absence of tumors in the controls can be questioned on the basis of historical experience, but the results in the concurrent controls cannot be completely ignored without a fully adequate study to provide further assurance that the control group was truly aberrant.

(iii) *Site Specificity/Target Organ.* It is true that most carcinogens cause cancer at a specific site. However, as the Bureau maintains (Reply, pp. 15–16) and the record shows, a substance can be considered a carcinogen even though it induces tumors at various sites. G–219, p. 7, G–227, pp. 22–23. The definition of "cancer" developed by an FDA advisory committee (G–61) does not require tumor site specificity. The "Shubick" subcommittee, which developed a report, entitled "General Criteria for Assessing the Evidence for Carcinogenicity of Chemical Substances" for the National Cancer Institute, adopted a similar position. G–B.

As CCMA points outs, FDA has stated, in the November 4, 1974 **Federal Register** (39 FR 28908), that the fact that tumors were of different kinds in a rat feeding study on acrylonitrile "suggest no relationship to the agent fed." The acrylonitrile study was inadequate for several reasons, though, and the ultimate FDA action at that point was to require further studies to resolve questions concerning the safety of the substance. Thus, the FDA action is not a precedent that a cancer finding should be ignored if the tumors do not have a specific site.

It is not unknown for carcinogens to cause tumors at various specific sites. Nitrosamines, are "a good example of chemicals that have more than one target organ," but "generally" carcinogens have a specific target organ. Newberne, Tr., pp. 920–21.

The absence of any target organ can warrant further investigation and studies before a firm conclusion is drawn about the carcinogenicity of a substance. However, I do not believe that site specificity should be regarded as an essential requirement before a substance is considered a carcinogen. Not enough is known about cancer to assume that it always has to be site specific. More importantly for the present case, I do not believe that positive cancer findings should be completely ignored because they lack a specific target organ. At the least, substantial question as to carcinogenicity remains and a further study is needed to resolve the question whether the substance is a carcinogen.

Lastly, the statistical findings of lymphosarcomas in female rats made by Schneiderman and Gaylor shows a site specific effect, although CCMA maintains that these lymphosarcomas are not sufficiently site specific. CCMA also argues the finding is invalid because it results from the use of different pathological screens. This objection is discussed in the next paragraph.

(iv) *Lymphosarcomas Analysis and Full Pathological review.* Dr. Schneiderman and Dr. Gaylor found a statistically significant dose-related linear trend for lymphosarcomas in female animals. Schneiderman, G–227, p. 18, Gaylor, G–226, p. 25.

The statistical analysis was done on the four feeding levels and control level involved in the Taylor/Monlux study, but only three levels had full pathological screens, and the other two levels had partial short screens. CCMA objects to any reliance on this statistical finding due to the different pathological screens. According to CCMA, Dr. Schneiderman recognized that using different screens presents "a real problem" and admitted that the trend would be "weak" if only the fully screened levels had been used. Tr., pp. 766, 834 as cited in EX., pp. 75–76. CCMA maintains that, if full screens had been done for the levels for which short screens were done, it is "likely" more tumors would have been found, and these findings would reduce the significance of the dose response trend that Drs. Gaylor and Schneiderman found.

In its reply, the Bureau points out that, if a certain number of additional tumors were found on a full screen in the levels partially screened, the trend would still be significant and linear. Moreover, if even more tumors were found in the partially screened middle levels, the trend would not be linear but would present a response pattern with a hump in the middle. Such a pattern would suggest that Red No. 2 induces more cancer at lower doses than has been believed. Reply, pp. 17–18.

I reject the CCMA position that the positive findings of lymphosarcomas in this study should be overlooked because the determination was made on the basis of different screens. CCMA urges that only the fully screened feeding levels be examined, and the trend for these levels disregarded as weak. Dr. Schneiderman in his testimony regarded the different pathological screens as a real problem but one that arises because:

> "anytime an experiment is done at less than the optimal way and the materials looked at not as fully as they should, what errors are named (sic) in this kind of thing, you will get less of an effect." Tr., p. 828.

Despite the drawback that full pathological screens were not available, Dr. Schneiderman thought an analysis for dose response could still be made on the basis of the available data from all the feeding levels. Tr., p. 827.

The lymphosarcoma trend was the only dose-response effect found in the study, and it is important because the presence of a dose-response effect makes scientists believe more strongly that a substance is a carcinogen. Schneiderman, G–227, p. 18. If a dose-response effect were not present for Red No. 2, but there were a significant increase in cancers in the treated animals with the controls, there would be less confidence that the substance is a carcinogen, but the results still could not simply be ignored. It would be appropriate at the least to require a full and adequate study to resolve the question whether there is a carcinogenic effect from the substance.

I believed that weight should be given to the positive finding even though it was not made on the basis of full screens and might have been different if more information had been available. I do not believe it is reasonable to consider Red No. 2 safe on the basis of hypothetical estimates of what the tumor findings could have been if full screens have been available, unless it is reasonably certain that no significant findings of cancer could have been made if all the information had been available. Instead, the matter is unclear; it is possible that there is a dose-response effect, or an unusual pattern of tumor findings. Thus, these results require additional study.

CCMA also disputes reliance on this finding of a dose-response effect because the trend was found only in female rats, and there is no explanation for why these tumors would occur in only one sex. However, as Dr. Gaylor reported, the male rats did not live as long as the females and did not have as extensive pathology as the female rats.

G–226, p. 25. this fact could account for the difference in effect. Moréover, it would be inappropriate to ignore an effect found only in one sex even though we do not know a reason why the effect is or should be sex-specific. A carcinogenic effect found in only one sex raises the question whether the substance is a carcinogen, even though until more information is available it is not possible to explain satisfactorily the limited scope of the effect.

I also agree with the Bureau that the statistically significant positive finding for lymphosarcomas deserves weight even though there are other analyses with negative findings for other tumor categories. Reply, pp. 18–19.

(v) *Total Malignant Tumor Trend in Female Rats.* Dr. Gaylor testified that there was a statistically significant increase at the .04 level in malignant tumor in the high dose female "terminal" animals (those that survived 30 months), as compared with the terminal control animials. G–226, p. 21. These controls had no tumors, an unusual finding. If "interim" animals (those that did not survive 30 months) as well as terminal female controls are examined, the tumor rate is similar to the rate in other control groups in this study. A comparison of the .03 percent dose level terminal and interim female rats with these controls shows an increase in malignant tumors of all types, but the increase has a weaker level of statistical significance of .07. G–226, pp. 23–24. The high dose group in a similar comparison exhibited a statistical significance of .08. Ibid.

As already discussed, CCMA objects to this analysis due to the lack of specific tumor sites and the low rate of tumors in the controls. CCMA also objects to any reliance on this finding because it was made using "Fisher's Exact Test," a test developed for fixed row and column figures. The Bureau maintains that this test can appropriately be used for other purposes, and cites statistical authorities in support of its position. Reply p. 20, Tr., p. 966.

I find the Bureau's reply convincing, and I also accept Dr. Gaylor's explanation that the Fisher Exact Test is appropriately used here due to the small number of animals involved in the analysis. G–226, p. 38. Moreover, it is appropriate to take into account Dr. Gaylor's experience in carcinogenicity testing since this matter involves a judgment about which statistical procedure provides useful information for evaluation. Dr. Gaylor is chief of biometrics at the National Center for Toxicological Research.

Dr. Gaylor's analysis found an increase in tumors for terminal females at the .04 level of significance. Although the tumor increase in terminal female animals was within the 5 percent limit for statistical significance commonly used in the scientific community, the tumor increase for interim and terminal animals was outside it. CCMA criticizes reliance on the findings that fall begond the 5 percent limit.

The interim and terminal animals were combined for analysis in recognition of the argument that the results for the terminal animals were questionable due to the unusual absence of tumors in the controls. As already discussed, it is not clear that the zero incidence in the controls should be disregarded. Furthermore, I agree with Dr. Gaylor that the 5 percent level of statistical significance is often used as a benchmark, but it should not be given absolute significance. G–226, p. 3; see Cyclamates, Notice of Interlocutory Decision, 44 FR 47620 (August 14, 1979), and Benylin, Notice of Final Decision, 44 FR 51520–51521 (August 31, 1979). The positive findings, even though at a lower level of significance, merit some weight as suggesting a possible carcinogenic effect for Red No. 2 on this record, in light of the more significant positive finding for the terminal animals and in light of the lymphosarcoma findings.

Moreover, if the original pathological findings as reported in the Taylor/Monlux study were used, the positive findings in Dr. Gaylor's analysis would be of stronger statistical significance and within the .05 limit. Instead, Dr. Gaylor used the revised pathological figures that were developed when FDA reviewed the data sometime later.

CCMA seems to argue that Dr. Gaylor's revised analysis should be given little weight because different statistical procedures were used after the revised pathological tumor findings reduced the significance of the original findings, which resulted from other test procedures and significance levels, and were based on the original pathological review. It is not clear whether CCMA is raising this point as an exception, but since it is referred to at several places in CCMA's exceptions, I will address the matter. Ex., pp. 16, 77–78.

It should be borne in mind that the TAC requested Dr. Gaylor to consider biostatistical aspects of the study. P–74, p. 18. Dr. Gaylor on cross-examination acknowledged that he changed the procedure he used for statistical analysis of the original figures and the revised figures. Tr., p. 744. He used the two-tail test and a .05 level as the mark of statistical significance for the original figures. If these procedures were applied to the revised figures, the results would not be statistically significant at the .05 level. Dr. Gaylor maintained, nonetheless, that the .05 level is only a benchmark, not an absolute standard, and that a one-tail test is the appropriate test, and probably should have been used initially. I agree that the one-tail test is appropriately used when seeking to find out whether a substance causes a harmful effect. A two-tail test halves the significance of findings, and is appropriately used if the purpose of the test is to compare treatments. Schneiderman, G–227, p. 19. The purpose of the statistical test in this case is to detect the ability of a substance to cause an adverse effect, and it is important to use test procedures that will provide assurance of detecting an effect of this type. Therefore, I conclude that it is appropriate to rely on Dr. Gaylor's revised analysis and the statistical procedures used.

Lastly, CCMA appears to be criticizing Dr. Gaylor's total tumor analysis, and the references in the Initial Decision to total tumor increase, because Dr. Schneiderman did not report the same analysis. Ex., pp. 86–90. Dr. Schneiderman testified, however, that "by and large we agreed with what Dr. Gaylor had found." G–227, p. 19. Furthermore, Dr. Schnederman reported that he found an increase in total tumors in female rats, which was "barely significant" at the .048 level. When this finding was "corrected * * * for the survivial" of animals, the trend was below the .05 benchmark of statistical significance. G–227, p. 17. Similar results were found by Dr. Gaylor, who reported an increase in total tumors for terminal female rats (significant at the .04 level), but a less significant trend for terminal and interim rats. G–226, pp. 37–38. As Dr. Schneiderman also testified when tumor types were examined, a statistically significant increase in lymphosarcomas in females was found, a result also reported by Dr. Gaylor. G–266 pp. 25–26.

I believe the lymphosarcoma finding is a key finding that is sufficient by itself to raise a serious question about the carcinogenicity of Red No. 2. The finding of a total tumor increase in female rats is also important because it reinforces the concerns about the carcinogenicity of Red No. 2 arising from the other finding. This total tumor finding is appropriately relied on as part of the overall basis for concluding that Red No. 2 has not been shown with reasonable certainty not to be a carcinogen.

(vi) *Conclusion.* I conclude that there is disturbing evidence of possible carcinogenic effects from Red No. 2 based on the lymphosarcoma finding, singly and in combination with the total

tumor increase in female rats in the Taylor/Monlux study.

B. *Andrianova (Russia) Study (G–14).*—1. *Introduction.* The Andrianova study (G–14) was conducted in the USSR to test the carcinogenic potential of several dyes, including "amaranth," an internationally used name for the substance called FD&C Red No. 2. The Russian investigators concluded that:

> Administration of chemically pure amaranth to rats during 33 months resulted in the development of 15 malignant tumors in 13 rats out of 48 surviving at the time when the first tumor appeared * * *. In the control group, where the rats did not receive any dye, tumors were not discovered * * *. Chemically pure amaranth possesses carcinogenic activity of medium strength and should not be used in food industry. G–14, p. 7.

However, questions have long existed about the purity and identity of the substance tested in the Andrianova study. In the FDA decision denying the petition for permanent listing of Red No. 2—the action that led to this hearing—FDA stated that "questions remain as to the chemical equivalence of FD&C Red No. 2 to the amaranth used by the Russian investigators." 41 FR 15053 (April 9, 1976).

FDA made several efforts to obtain more information about the substance tested in the Andrianova study. These efforts, which proved to be unsuccessful, were described to the TAC as follows:

> With respect to the second Russian study which showed positive carcinogenic effects, Dr. Murphy asked if the Agency had approached the Russians to determine the composition of the material that they used. Mr. Gittes replied that in 1971, the FDA asked the State Department to find out the composition of the amaranth material used by the Russians and the spontaneous rate of tumor occurrence in the Russian rat. However, the FDA has received no response from the State Department to date on these questions. Other attempts through various Russian investigators, FDA personnel traveling to Russia, etc., to obtain answers to these questions have been similarly unsuccessful. P–74, pp. 3–4.

2. *Initial Decision.* The Initial Decision found that the Andrianova study (G–14), was conducted "to determine the tumorigenic effects of Red No. 2 on rats," and that "[p]ositive results were * * * obtained in this study." ID, p. 16. The Appendix states in connection with this study that "[a]bsent overriding factors, substantial weight" should be given to the conclusion of the researchers about their own study. ID App. p. 35.

The Appendix to the Initial Decision apparently rejected the CCMA argument that the Red No. 2 used in this test was impure and not the same as FDA-certified FD&C Red No. 2 because the papers suggesting the presence of an impurity were "hearsay" and "there is no evidence of record to indicate that the substance tested was anything other than chemically pure (Red 2)." ID App., p. 34. CCMA's objections based on the unusual factor that the controls had a zero tumor incidence rate were not given weight because spontaneous rates of tumors vary among test animals and suppositions about the expected rate cannot be based on historical patterns for a different strain of rat. ID App., p. 34.

3. *Exceptions.* CCMA believes that the ALJ was "completely in error in giving the slightest credence to the Andrianova study, which does not provide any evidence showing FD&C Red No. 2 could be carcinogenic." Ex., p. 109.

Memoranda in the record show that an FDA scientist, Dr. Herbert Blumenthal, who attended an international meeting held to "shed some light on the type of amaranth that was used," reported that there was "substantive evidence" to show that the color was not FDA certifiable Red No. 2. Ex., pp. 4, 5–6, citing P–145. The International Association for Research in Cancer has rejected the Andrianova study (P–145), as has the FAO/WHO, because the study used "samples of amaranth" with different specifications. P–16, pp. 13–14. The CCMA also maintains that there is "doubt in the scientific community" about the existence of controls in the study or that they actually had a zero incidence level of tumors." Ex., p. 108.

4. *Bureau Reply.* The Bureau argues that the uncross-examined written reports that the substance tested by the Russians may not have been Red No. 2 should be given "very little weight, if not totally ignored." Reply p. 25. In addition, it maintains that even if the amaranth used by the Russians differed from the FDA color no one knows whether the difference would affect the results in the carcinogenicity testing. Consequently CCMA has not met its burden to prove Red No. 2 safe.

5. *Analysis.*—(a) *Chemical Equivalence.* I reject the position in the Initial Decision and the Bureau reply that the criticism of the Russian study should be given little weight or ignored simply because the criticisms are found in written reports and have not been subject to cross-examination. FDA has long taken the position that the positive results in this Russian study do not establish that Red No. 2 is a carcinogen because of questions about "the chemical equivalence" of FD&C Red No. 2 to the amaranth used by the Russian investigators. 41 FR 15053 (April 9, 1976). The written material should be credited on the basis of what it reports, taking into account any factors that reduce its reliability.

In reviewing the written material cited in the exceptions as bearing on chemical equivalence, it should be noted that the questions relate to whether the amaranth used by the Russians was as chemically pure as FDA certified Red No. 2. The writers of these memoranda believe the Russians may have used a less pure version of Red No. 2 and not a totally different substance, as is evident from the quotations given above from the documents cited by the petitioner in its exceptions. Consequently, the fact that positive results were found creates a significant question as to whether the results are attributable to the impurities that may have been present in the Russian sample, or to FD&C Red No. 2 in its FDA-certifiable form. The material cited by CCMA does not provide definitive information to show that the results are attributable to impurities that may not occur in FDA-certified Red No. 2.

Thus, I find that due to the question about the chemical equivalence of the test substance to FDA certified FD&C Red No. 2, the positive findings in this study do not dispositively establish that Red No. 2 is a carcinogen. The study cannot be ignored entirely, however. Instead, it creates a concern that Red No. 2 may be a carcinogen. This concern needs to be resolved by an adequate test before Red No. 2 can be approved as safe for use.

(b) *Controls and General Unreliability.* CCMA has also urged that this study should be entirely ignored on the ground that it has been rejected by international groups such as FAO–WHO, because there is "doubt in the scientific community" about the zero incidence level of tumors reported for the controls in this study. Ex., pp. 108, citing, for example, P–157 at 27, P–11, P–16, pp. 13–15, P–62, P–65, P–68.

The "most important" question about this study concerns the chemical identity of the test substance, the matter already discussed, according to the memoranda cited in the exceptions (P–145). It is true that the FAO–WHO did not fully accept this study, and it retained a "temporary" acceptable daily intake level for Red No. 2, but it also expressed the hope that "international cooperative studies with standard samples may be undertaken to elucidate the reason for the reported discrepancies." P–16, pp.13–14.

I agree that the zero incidence level of tumors in the negative controls in the Andrianova study is unusual and gives rise to some questions, as discussed in

connection with the Taylor/Monlux study. A zero incidence level is possible, though, due to variations in spontaneous tumors.

Moreover, in the Andrianova study other red dyes were also tested, and amaranth had a high, statisically significant increase in malignancies at the 2 percent level compared with rats fed another dye, according to Dr. Gaylor's analysis of the results. G–226, p. 29. Those rats fed the other dye had tumors, and thus did not have the unusual absence of tumors reported for the negative controls. Consequently, the positive results observed from amaranth in comparison with a positive control group that had tumors tends to obviate any questions about the significance of the effects of amaranth compared with negative controls that had no tumors.

Nonetheless, there continues to be a question whether the design of the study meets current standards for a fully adequate study in view of its limitation to male rats. A variety of tumors were reported, and as discussed in connection with the Taylor/Monlux study, the presence of a site specific response increases the confidence that a substance is a carcinogen. But despite questions of this type a study with positive findings of carcinogenicity cannot be ignored in the absence of a completely reliable study that shows with reasonable assurance that the substance is not carcinogenic.

C. *Other Chronic Studies; Webb Rat (P–2), Mouse (P–3) and Dog (P–5) Studies.*—1. *Initial Decision.* The Initial Decision includes the following discussion about the Webb rat, mouse, and dog studies conducted by FDA:

> Several studies have been conducted in mammals in order to directly observe any tumoringenic tendencies of Red No. 2. Dr. Webb and his associates completed the studies on the dog, mouse, and rat in which the animal ingested Red No. 2 as some percentage of its food supply. The mouse and rat studies were very similar in test protocol and implementation. Two strains of animals were used with heavy mortality in one of the strains in each of the two tests. Limited histopathological examination was performed. In general, this examination was limited to those tissues which, on gross examination, appeared to contain a tumor or other lesion.
>
> In the Webb rat study, the Osborne-Mendel rats did not show tumor development associated with the ingestion of Red No. 2. However, among the Sprague-Dawley rats, there was a doubling of breast tumors in the treated animals as compared to the controls.
>
> In the Webb mouse study, there was an increase in lymphosarcomas in the treated as compared to the control animals. There was also a general increase in tumors in the treated male mice as compared to the control male mice. These two studies raise questions as to the carcinogenicity of Red No. 2 because positive results were partially demonstrated.
>
> Dr. Webb also conducted a study in five dogs to determine if the ingestion of Red No. 2 produced tumors. None were seen in this study. However, the dog is generally considered as being useful in revealing human carcinogens only in the case of bladder tumors. ID pp. 15–16.

2. *Webb Rat Study (Osborne-Mendel Strain).*—(a) *Exceptions.* CCMA maintains that the Webb rat study of the Osborne-Mendel strain of rats support the non-carcinogenicity of Red No. 2. It points out that the FDA scientists who did the studies regarded the results as negative. Ex., pp. 90–166. It disputes the Bureau's position that the histopathological examination in the rat study was inadequate on the ground that FDA scientists and advisory committees regarded the study as adequate for many years. CCMA maintains that the likelihood of "overlooking any important tumors was 'fairly low,'" despite the "limited histopathology," because tumors can usually be detected by gross examination. Ex., pp. 96–97. One of the petitioner's witnesses who helped develop the NCI guidelines testified that the "standard protocol" requires a microscopic examination of all the tissues from the high dose and control groups, and a microscopic look "only at grossly altered tissue and target organs from the low and intermediate levels." Busey, P. 152, p. 12. CCMA maintains that this type of examination is "precisely" what was done in the Webb study. Ex., p. 96.

(b) *Reply.* The Bureau in its reply maintains that current practice and the NCI guidelines with respect to histopathological examination were not followed. Reply, p. 22. The Bureau also regards the Webb rat study as deficient due to defects in gross pathology, the failure of CCMA to show Dr. Webb's qualifications, and other factors. Brief, pp. 12–16.

(c) *Analysis.*—(i) *Histopathological Examination.* In this study, "[a]bout 220 pieces of tissues" were sectioned and microscopically examined, "including all malignant and doubtful appearing masses, and various other lesions." P–2, p. 1. Under the NCI guidelines, a histopathological examination is to be made of all treated and control animals in chronic studies and the examination is to include over 20 specified tissue sites in addition to gross lesions and suspect tumors. G–11, p. 55. The Bureau maintains that, under current standards, full screens, under a microscope, are required of the high dose animals and controls, or at least a significant proportion of them, and a partial screen of tissues is needed of all the remaining animals. Reply, pp. 22–23. CCMA's pathology witness, Dr. Busey, also stated that full screens are needed for the high-dose and control groups, P–152, p. 12, see P–150, p. 35. However, in this study, no full screens were done of any feeding levels.

The Director of the Bureau's Division of Pathology (Weinberger G–218, p. 14) testified that "in some of the animals was a detailed or 'full screen' examination of tissues undertaken," and that "only a very small number of tissues" were examined. G–218, p. 14. In his opinion, the examination of "200 tissues" from 870 rats was "markedly deficient," and this methodology made it "almost impossible" to pick up small neoplasms, and "very difficult" to determine the primary site of others. Id. Because of these and other defects in this study, he doubted that the study could detect a weak carcinogen or subtle toxic effects, and had "strong reservations" about accepting its negative conclusions. G–218, p. 16.

At the intermediate dose feeding levels, CCMA believes that under the standard procotol it is sufficient to have a histopathologic examination of only "grossly altered tissues and target organs." P–152, p. 12. The Bureau believes that at least "partial screens" are necessary of 5–9 tissues from each remaining animal. Reply, pp. 22–23; Brief, p. 13. The NCI guidelines require an examination of a number of tissues in addition to those with obvious or suspect tumors. G–11, p. 55.

CCMA maintains that "representative samples" were examined in the Webb study in addition to abnormal tissues. Ex., p. 92. However, the sampling of "various other lesions" in the Webb study cannot be regarded as sufficient to satisfy current guidelines in view of the enormous difference between the number of tissues examined in the study (220 tissues) and the number examined in partial screening (4,000–7,000 tissues) of a full screening (11,000 tissues) in a study complying with NCI guidelines. Bureau Brief, p. 13 citing Weinberger. I find CCMA's position on whether the Webb study satisfies the NCI guidelines ambiguous. At some points CCMA states that the histopathology was "limited" and less than full (Ex., pp. 92, 96), but at another place it states that the NCI guidelines are "precisely what was done" in this study (Ex., p. 96). This description, however, may relate to a claim of compliance with the type of partial screening needed at intermediate feeding levels, rather than to the type of

full screening needed at high dose and control levels.

In any event, for the reasons discussed above, I find that the Webb rat study does not meet the current standards for histopathologic examinations as found in the NCI guidelines because of the failure to have any full screens, and the failure to have at least partial screens of all animals.

CCMA also argues that the histopathological evaluation can be adequate even when the histopathology is limited because most tumors can be spotted by gross examination. Petitioner's witnesses estimated that "about 90%," "most neoplasms," and "about 95%" as "a rough figure" of all tumors would be detected by gross examination, for example, P–150(2), p. 35, P–152(2), pp. 12–13; P–156, pp. 16–17. On the other hand, a Bureau witness estimated that about 20 percent of tumors could be missed if only grossly visible neoplasms are given a microscopic examination. G–218, pp. 6–7. There were no published studies to show the number of tumors that would be missed, but the Bureau's witness cited examples of misses of 12 percent and one of 27 percent. Weinberger, Tr. pp. 560–61. Moreover it is clear that leukemias would be missed without a microscopic examination. Ex., p. 100. Lymphosarcomas are harder to diagnose without a microscope. Newberne P. 161, p. 31. Thus, it is clear that the failure to meet the current standards for histopathology reduced the ability of this test to detect any carcinogenic potential of Red No. 2. This deficiency, while disturbing at any time, makes it especially inappropriate to rely on this study as showing the noncarcinogenicity of Red No. 2 in view of the other indications of possible carcinogenicity of the color found in the later Andrianova and Taylor/Monlux studies. See also ID App., p. 19.

(ii) *Pathology Work.* Furthermore, the pathology would have to be of high quality if test results were to be relied on despite a histopathological review of inordinately limited scope. Here, there is reason for regarding the pathology work as of questionable quality.

The gross pathology was done on shrunken, hardened and discolored tissues fixed in formalin, rather than on freshly killed tissue. G–218, p. 14. CCMA witnesses did not think this created a great deal of difference and stated that FDA currently has tissues examined in formalin. P–153, p. 17, P–161, pp. 12–13. A Bureau witness, however, believed that the tissue-fixing methods used in the Webb study made observations "of necessity limited." G–218, p. 14. Use of preserved tissues is not a preferred practice. Tr., p. 384. Thus, I find that the use of tissues of this type, in the absence of histopathological examination, reduces confidence that the gross examination was adequate.

Moreover, there were several differences between the diagnosis made in the study and the diagnosis made by the petitioner's pathology expert (Busey, P–152(2)) in a review of the tissues in preparation for the hearing. The Bureau's pathologist believed that a comparison of individual animal diagnoses of Drs. Busey and Webb showed essential agreement in 50 percent, and "substantial differences" in 50 percent Weinberger, G–272, p. 7. The petitioner's pathology expert thought the differences minor (P–152(2), p. 14), and another of its experts believed the differences involved "less than 18 percent" if the calculations were made on the basis of all the tissues actually examined, and if "similar" diagnoses were eliminated. Newberne, P–161, p. 6. Cross-examination showed, however, that there is some basis for regarding the "similar" diagnosis as having important differences. Tr., pp. 890–93.

Thus, more than 18 percent of the diagnosis could be considered to be of questionable quality. The Bureau also believes that a calculation based on differences in positive diagnoses in animals provides a better reflection of the quality of the pathology work than differences in the diagnoses of all slides, because a diagnosis of no tumors is easier to make than the classification of a tumor which is found. Bureau Brief, pp. 14–15. Consequently, there is a substantial basis for regarding the pathology work as being of questionable quality. It would be inappropriate to assume that the gross pathology was so well done that it can be relied on to provide a basis for estimating the tumor incidence in the Webb studies even though the histopathological review was limited.

I find it unnecessary to reach the issue of whether the Webb study should be disregarded as deficient solely because CCMA did not show Dr. Webb's qualifications as a pathologist. The Bureau maintains that CCMA has the burden of proof to show the qualifications of Dr. Webb because CCMA seeks to rely on the study done by him. Bureau Brief, p. 22. According to the Bureau, Dr. Webb was a Doctor of Veterinary Medicine (DVM), and the petitioner's witnesses recognize that DVM training is insufficient by itself to qualify an individual as a pathologist. Bureau Brief, p. 22 citing Radomski, Tr., p. 254, Busey, Tr., p. 351. On ther other hand, CCMA urges that "a pathologist qualified to be employed" by FDA should be assumed to be qualified, unless FDA affirmatively shows his lack of qualifications. Ex., p. 102. Although I find the record unclear on whether Dr. Webb was specifically employed by FDA as a pathologist, it is clear he was assigned to do pathology work in the Webb study. I agree that an individual assigned to do pathology work by the Bureau should be assumed to be qualified for the assignment, unless the Bureau shows his lack of qualifications. However, even assuming Dr. Webb's qualifications for the initial assignment, the Webb study cannot be considered adequate to show the safety of Red No. 2, due to the insufficient amount of histopathology, and the questionable quality of the diagnoses actually made in the pathological evaluation, as discussed above.

(iii) *Limitations of Detection under NCI Protocol.* CCMA also seems to maintain that the Webb study should be considered adequate because the Bureau's pathologist testified that the test could detect a potent carcinogen, but not a weak one, and the same results occur under the NCI's criteria because all that can be detected under the standard NCI protocoal is a potent carcinogen. Ex., p. 101–02, citing Gaylor, Tr., p. 780; Newberne P. 161, pp. 6–7.

As Dr. Weinberger testified (G–218, p. 13), the Webb study could probably have detected a potent carcinogen, but he had "strong reservations" about its ability to detect "a moderate or especially a weak carcinogen." G–218, p. 13. It is also true that bioassays conforming to the NCI guidelines are most able to detect potent carcinogens and have less ability to detect weak carcinogens. But this limitation on the capabilities of tests meeting the NCI guidelines does not mean that the Webb study had an equivalent capability to detect a carcinogenic effect as a study fully complying with the NCI guidelines on histopathology. Dr. Weinberger testified as follows on cross-examination in response to questioning on this matter by CCMA (Tr., pp. 570–71):

> [Answer] I think that it is hard to be absolute in these answers, but I think that if the NCI guidelines were followed exactly as indicated, the possibility of picking up a week carcinogen would be much better, greatly augmented, over this study. I am not saying it would pick it up in every situation. You have to qualify it, how weak and what percentage and so on.
>
> [Question] Let us just try, if we may, again. Pursuant to the NCI guidelines, if one conducted a carcinogenesis study, would it detect, considering what those quidelines are, would it detect a weak carcinogen?

[Answer] I think in many cases it would.

* * * * *

[Question] So you testimony is that a weak carcinogen would be detected—

[Answer] Yes, at 5 percent.

[Question]—by a carcinogenesis study—

[Answer] I do not say it always would, but I say it certainly could.

[Question] How often, Doctor, in your opninion? Do you have an opinion?

[Answer] I could not say. I have no opinion.

Dr. Schneiderman also testified that some NCI studies suffer from the defect seen in the Webb and other early FDA studies, "but not nearly as much." G–227, p. 16.

Thus, the difficulty in detecting weak carcinogens makes it important to have tests that are as adequate as possible. The Webb study was clearly not as adequate as tests meeting current guidelines.

(iv) *Conclusion as to Webb Rat Study.* I believe that the deficiencies shown by the Bureau on Webb Rat Study, notably those concerning the inadequacy of the histopathology and gross pathology, make it inappropriate to rely on this study as showing to a reasonable certainty that Red No. 2 is not a carcinogen. I accept the Initial Decision to the extent it makes this finding.

3. *Evaluation of Sprague-Dawley Strain of Rats.* In the Webb study, two strains of rats were tested. One of the strains (Osborne-Mendel) had an autolysis rate that was acceptable and a survival rate that was borderline-acceptable (Weinberger, G–218, p. 15), but the study of this strain of rats was inadequate due to the deficiencies discussed above. The study of the other strain of rats (Sprague-Dawley) had other deficiencies. CCMA concedes that "too few" of the Sprague-Dawley rats survived to permit "meaningful analysis." Ex., p. 104. Consequently, CCMA believes that the Initial Decision was incorrect in giving any significance to the doubling of the breast tumor rate seen in the Sprague-Dawley strain. Ex., p. 105.

CCMA excepts to the ALJ's "efforts to dream up" some positive findings in the Webb study by focusing on this increase in breast tumors in the Sprague-Dawley rats. Ex., p. 105. According to CCMA, the survival rate was not adequate in this strain to permit statistical analysis, the increase was not significant, and a Bureau's pathology witness (G–217, p. 14) did not find the tumors significant. Ex. p. 105–06.

A poor survival rate makes it more difficult to detect an effect. It is arguable, therefore, that the finding as to breast tumors is entitled to some weight. G–226, pp. 13–14. Nevertheless, I do not believe that the observation of breast tumors in this strain of rats should be relied on as a part of the basis in this record for suspecting the carcinogenicity of Red No. 2 due to the lack of significance attributed to the study and the type of tumor by a Bureau witness (G–217, pp. 14–15), and the fact that in its brief the Bureau made only passing references to the finding without any further explanation of the Bureau's position. Brief, p. 24.

4. *Webb Mouse Study.* CCMA discussed the Webb rat study in its exceptions, but did not discuss the Webb mouse study separately. According to CCMA, the studies "were conducted under the same protocols by the same scientists; produced the same negative results; and have been subject to the same criticism. Accordingly, we here discuss only the rat study, since the same views apply to the mouse study." Ex., p. 91, n. 53.

I agree and find no need for an extensive separate analysis. The Webb rat study and the Webb mouse study presented similar issues, as the petitioner has pointed out. The evaluation of the exceptions concerning the Webb rat study has provided an adequate opportunity to evaluate the significance of the Webb mouse study, and to conclude that the latter does not establish the safety of Red No. 2.

The findings of fact concerning the Webb mouse study proposed by the Bureau in its brief to the Administrative Law Judge have been reviewed and adopted with one change. I have omitted to make any finding concerning the presence of a suggestive increase in male mouse tumors in the Webb mouse study because it is unnecessary to reach that issue. The other findings adopted in this decision adequately support the conclusion that the Webb mouse study does not demonstrate the safety of Red No. 2.

5. *Dog Study (P–5).* CCMA in its exceptions describes the Fitzhugh/Kasza dog study as providing further evidence that Red No. 2 is not a carcinogen, Ex., p. 90.

The Initial Decision contained an extensive discussion of the parties' positions on this study, and a description of the limitations of the study. The study's usefulness is primarily limited to helping evaluate the risks of bladder cancer. ID App., p. 23–26. The Bureau's reply emphasizes that a 7-year dog study is not a lifetime study (G–217, p. 9, Tr., p. 901), and, despite use of studies of this length for economic reasons, longer studies should be used when needed to make an adequate toxicological evaluation. Reply, p. 27. Moreover, the study had too few animals to be an adequate study. G–226, p. 30, G–218, p. 18, G–217, p. 8. A dog study of this length in these circumstances does not provide a sufficient basis to prove safety. Accordingly, I agree with the Initial Decision Appendix that this study is not useful "as a general finding of safety." ID App., p. 26.

D. *Short-Term Screening Studies.*—1. *Introduction.* Several short-term screening studies were conducted on Red No. 2. These studies attempted to detect the ability of Red No. 2 to induce mutagenic changes in cells because mutagenesis "has been shown, to some extent, to be predictive of a chemical's carcinogenic potential in higher animals." ID, p. 12. Positive mutagenic responses were observed with Red No. 2 or its known metabolites in one or more tests in the following test systems: cytogenic, Oster Drosophilia, mouse lymphoma and dominant lethal. The Ames test and yeast studies were negative. The host mediated assay was apparently negative.

2. *Initial Decision.* The Initial Decision provided the following summary of the short-term screening tests conducted, and their results.

> One type of microbial mutagenesis test run on Red No. 2 was the Ames test. Two such studies showed that Red No. 2 was not a mutagen in this test system. The mutagenic effects of Red No. 2 on yeast were also analyzed. Two of these tests were conducted and while one initially indicated positive results, the researchers attributed this to a "contaminant." Ultimately, the tests showed negative results.
>
> The host mediated assay is designed to test the mutagenicity of microorganisms exposed to the chemical through the ingestion of the chemical by an animal which contains the microorganism in its peritoneal cavity. This test method more closely approximates the range of metabolites which occur in the whole animal system. Two tests were conducted by Stanford Research Institute (SRI). While positive results were obtained initially, this effect was attributed to a contaminant and the second test showed no positive results.
>
> Mouse lymphoma tissue culture cells were also used to test for any mutagenic effect of Red No. 2. The chemical was positive for mutagenicity in this test system, and demonstrated a dose response. A series of tests on higher animals was also conducted to determine if there was any mutagenic effect from exposure to Red No. 2.
>
> Three studies were conducted using rats to test for mutagenesis by the dominant lethal technique. The FDA study demonstrated positive results. The SRI study showed pre-implantation loss at high-dose levels. The Kennedy study showed the same effect as the SRI study at all dosing levels.
>
> Three Drosophila studies were conducted, each using different strains of the species. One of these showed a positive response, while the other showed negative results.

Mutagenicity of Red No. 2 was also tested by noting chromosomal damage to the bone marrow cells of treated animals in two studies. Positive results were obtained in the FDA study for Red No. 2 and amino-R-salt, while the SRI study did not show this effect. The SRI also conducted an anaphase study on human lung tissue. Mutagenic activity was noted in this study. ID, pp. 13–14.

The ALJ concluded:

The mutagenicity tests of Red No. 2 have culminated in mixed results. Some systems show consistently negative results, others show positive results, and a third class demonstrated results which varied from research group to research group. The apparently conflicting studies which used only a single test system often showed other differences, such as in the strain of the animal tested or a difference in the testing protocol. With such a variation in the results it is clear that the question of mutagenicity of Red No. 2 has not been definitively resolved. ID, p. 14.

3. *CCMA Exceptions.* In its original brief to the ALJ, CCMA argued that the short term tests showed Red No. 2 to lack carcinogenic potential. CCMA Brief, pp. 92, 137. In its exceptions, CCMA maintains solely that the tests, even if positive, only call for a chronic study, do not provide definitive answers, and cannot overcome the results in chronic feeding studies, which CCMA regards as negative. Ex., pp. 110–32. On this basis, CCMA excepts to the Initial Decision "insofar as it appears to find that the short term tests * * * raise serious questions with regard to possible carcinogenicity." Ex., p. 110.

CCMA also views the short-term screening studies as of "comparatively minor significance" and states that the Commissioner did not rely on these studies as the basis for denying the petition to list Red No. 2. Ex., p. 110. According to CCMA, greater weight be given to the chronic feeding studies, which in CCMA's view establish the noncarcinogenicity of the color. Ex., p. 23.

CCMA's exceptions contain 20 pages of discussion of the tests. In what seems to be the only summary of the basis for its objections, CCMA maintains that the mouse lymphoma test is unverified, and the "few" other positives are an "artifact of the testing procedure" or are "overcome" by the "more extensive negative tests" in the same system. Ex., p. 112.

CCMA also believes that ALJ distorted the record by viewing the short-term screening tests as relevant to mutagenicity when they should be viewed solely as a guide to carcinogenicity. Ex., p. 24.

4. *Bureau Reply.* The Bureau maintains that the positive findings show a "potential hazard" with respect to carcinogencity. The reply extensively discusses the CCMA's exceptions. Reply, pp. 28–41. The Bureau maintains that CCMA overemphasizes the negative Ames and yeast studies, which are not reasonably expected to be sensitive for Red No. 2, and that CCMA underestimates the significance of the positive results seen in some of the other tests.

5. *Analysis.* As the Bureau notes, the "mutagenicity of Red No. 2 is an issue in this proceeding only as the capacity of a chemical to cause mutations also demonstrates that it is, or may be, a carcinogen." Reply, p. 28, n. 15. I do not accept the Initial Decision to the extent it is interpreted as finding that mutagenicity is an independent basis for denying the petition since the parties agree that this basis for action is not an issue in this proceeding.

It is also clear that due to their limitations the short term screening studies cannot provide evidence to a reasonable certainty that Red No. 2 is not a carcinogen. CCMA does not assert in its exceptions that they could support such a finding even though it did maintain this position in its brief to the ALJ. CCMA now maintains simply that the studies are not definitive, even if positive, and cannot overcome the results of the chronic feeding studies, which CCMA views as negative. Ex., pp. 110, 112. Because I do not believe the chronic feeding studies done so far are fully adequate, the important issue is whether the positive findings in the short-term screening tests provide additional support for believing that a further chronic test is needed before the safety of Red No. 2 can be adequately evaluated. I believe that they do reinforce the conclusion from the review of the chronic feeding studies that the noncarcinogenicity of Red No. 2 has not been adequately shown.

It is unnecessary to discuss all the tests and exceptions in detail in view of the extensive analysis provided in the Initial Decision. ID App., pp. 1–17.

Instead, I will focus on the significance of the negative results in the Ames test, and the mixed results in the Drosophila tests on fruit flys the two test systems CCMA regards as having the most significance for evaluating the carcinogenic potential of Red No. 2, and the weight to be given overall to the results in the short-term screening study. Ex., p. 115.

(a) *Ames test.* The Ames test is rightly regarded as highly predictive of carcinogenic potential on this record with most experts considering it 90 percent or more reliable (G–224, P–151), but some believing it less reliable (G–223, p. 26, G–167, P–150(2), PP. 42–43.) The Ames tests on Red No. 2 were negative. As discussed in connection with CCMA's "general areas of exceptions," the Ames test is not a conclusive indicator of carcinogenic potential, and negative findings do not establish that a substance is safe. Dr. McCann of the Ames laboratory testified that negative results are not dispositive, despite the 90 predictive value of the tests observed in the experiments so far conducted:

One can never be sure that a negative response is a true negative, or whether it is one of the few kinds of carcinogens that the test doesn't detect very well. * * * I think in general a positive result carries more weight with it than a negative result. G–224, p. 6.

Furthermore, "complicating factors" reduce the reliability of the Ames test in predicting the carcinogenic potential of a substance like Red No. 2. Dr. McCann testified:

But in my personal opinion in trying to draw a conclusion for Red 2, I would feel much more confident in the negative after we have had a chance to test a number of the other azo dyes that have the larger ring structure and that do have good animal cancer data. G–224, p. 7.

Moreover, the metabolism of Red No. 2 in the gut, rather than in the liver, makes the Ames test less reliable. Dr McCann stated that the test "most likely does not perform" the gut-type reaction very well and that she thought the system would not detect metabolic changes that occur from this type of reaction. G–224, p. 8, Tr., p. 412.

CCMA, though, points out that the known metabolites of Red No. 2 have been tested in the Ames test and found negative. Ex., p. 116, citing P–90, p. 9, P–101. On this basis CCMA argues that the Ames test findings on the breakdown products of Red No. 2 provide assurance about the safety of the substance and its breakdonwn products as metabolized in the human gut. However, this argument does not provide assurance about the safety of the unknown metabolites of Red No. 2, which are discussed below.

Even though I have not generally resolved the "chemistry" issue in this proceeding, the metabolism of Red No. 2 is relevant to the evaluation of the mutagenicity testing of Red No. 2's metabolites. The record shows that 69 percent of Red No. 2's metabolites are fully identified, but that the other part has not been recovered and is unknown. G–68. The petitioner's expert thought that the unidentified metabolites of the R-amino-salt became irreversibly bound to other compounds and, although the breakdown products could not be recovered by him, they would be rapidly excreted and be harmless. P–157, pp. 32–

33. However, it is clear that at present nearly one-third of the metabolic breakdown products are unidentified. This lack of information precludes reliance on negative findings on the Ames tests conducted on Red No. 2 breakdown products because not all of the breakdown products of metabolism have been identified or tested. Moreover, as the Bureau suggests, the unaccounted for portion of the metabolic by-product R-amino-salt, whose breakdown is not fully understood, has only 2 sulfonic groups in it. If the R-amino-salt breaks down in more than 2 parts, some of these parts may not be sulfonated. Bureau Brief, p. 97. Accordingly, these breakdown products may not be as rapidly excreted and could be more potentially harmful than the other breakdown products of Red No. 2. Bureau Brief, pp. 96–97. Although this possibility is hypothetical, the possibility does underline the lack of reasonable certainty about the full identity of the breakdown products. Accordingly, it is inappropriate to rely on Ames tests done on the breakdown products of Red No. 2 when it is not fully established what the breakdown produts actually are when Red No. 2 is metabolized in the body.

CCMA suggests that the Bureau's witness, Dr. McCann, believed that the negative Ames tests on breakdown products of Red No. 2 were relevant because the metabolic by-products of Red No. 2 "were actually tested in the Ames system and found negative," Ex., p. 125, citing G–224, p. 8. Dr. McCann did testify that the breakdown products of Red No. 2 had been tested. However, it is clear from her full testimony that she testified about her expertise in the Ames test system, and that she did not claim to have expert knowledge about the metabolism of Red No. 2. G–224. It is also clear from the following testimony that she did not know whether all the metabolites of Red No. 2 have been found:

> Of course, there may be other metabolites that are generated in animals. And what proportion of those occur in the liver and what portion occur in the gut, I don't know. G–224, p. 8.

Thus CCMA is incorrect in suggesting that Dr. McCann's testimony establishes that there are negatives Ames test findings for all the actual metabolic by-products of Red No. 2.

(b) *Drosophila Tests.* As CCMA emphasizes in its exceptions, three laboratories conducted Drosophila tests and the results were negative in two of the tests and positive in one. According to CCMA the positive finding was "weak," the test involved many less fruit flies than the other tests and the controls in the positive test had an abnormally low rate of changes; therefore, the positive finding should not outweigh the negative findings in the same test series. Ex., pp. 118–21.

However, the positive result was seen in the "sex-linked lethal assay" (G–4) and this Drosophila test system is considerably more sensitive than the tests in which negative results were reported. Mayer, Tr. pp. 667–68, G–185, pp. 128–29, G–29, p. 277, G–28, p. 241, G–55, Brusick, Tr. 434. The smaller number of flies used in the positive test does not undercut its significance because the statistical evaluation took into account the smaller number of flies used and, nonetheless, the positive findings were statistically significant at the P=.02 level on first analysis, and above the P=.03 level in a later analysis. G–4, G–226, p. 45, Tr. pp. 415, 433.

The CCMA argument that the positive results should be discounted on the basis of a comparison with historical controls is not persuasive. The laboratory strain of flies used in the positive study had an historically low spontaneous rate, and the rate was consistent with those in the Red No. 2 experiment. Brusick Tr., p. 442; Mayer, Tr., p. 675. CCMA argues that the general historical experience in Drosophila should be used for comparison rather than the experience in one laboratory because all the flies belong to the same species, regardless of the strain. Ex., p. 120. The record indicates, however, that there are differences in strains. G–222, p. 28; see P–134, P–135.

Given the conflicts in results between positive and negative results in the same test system, I find that the negative results cannot outweigh the positive. More testing is needed to resolve the conflict, but in the absence of further testing with clear negative results, a question exists about the potential mutagenicity, and thus the potential carcinogenicity, of Red No. 2 on the basis of the positive results in the Drosophila test. G–222, p. 28.

(c) *Other Short-Term Screening Studies.* Positive findings were also reported in the cytogenetic, mouse lymphoma and dominant lethal tests. ID Ap., p. 5–10, 13–17. CCMA suggests that these positive findings in short-term screening studies are outweighed by the more numerous negative findings. Ex., pp. 112, 124. It would be inappropriate, however, to allow the positive findings in the Drosophila and other tests to be outweighed by the negative findings in the Ames test. As already discussed, the reliability of the Ames test for a substance like Red No. 2 is limited. The yeast test systems, which had negative results, may also be insensitive for a substance like Red No. 2 that is metabolized in the gut. G–222, p. 32. CCMA maintains that the host mediated study, which apparently had negative results, is sensitive for substances distributed to the "peritoneal cavity, i.e. gut" Ex., p. 126. However, the gut and peritoneal cavity are different. Reply, p. 33. Moreover, the insensitivity of this test is recognized by a CCMA witness (P–151, p. 26).

Furthermore, short-term screening studies have a limited sensitivity, and a chemical that produces a positive effect must be quite potent. Legator, G–223, pp. 19–20, Tr., 787. The positive finding in the cytogenetic test is particularly important because the test tends to be insensitive. Legator, G–223, pp. 20–21. The relative importance of the positive effects in the in vivo cytogenetic experiments is attested to by Dr. Legator, G–223, pp. 20–21:

> If I achieved negative results with every other test that one can do for mutagenicity, and if I had the cytogenetic evidence, I would say that that would be enough to declare a compound as a potential mutagen. In fact, probably the only criticism that can be made of the cytogenetic procedure is that it is likely to be insensitive. When you find a positive, that is sufficient to declare the compound as mutagenic. If I had a dozen or any number of in vitro tests that showed that the compound was or was not mutagenic, I would not list those in the same category as one good animal test such as the cytogenetic procedure.

Even though CCMA views the intraperitoneal route of administration as inappropriate in the cytogenetic test (Ex., p. 123, P–159), this route maximizes the effect, and is suitable for a short-term screening study of limited sensitivity. Tr., p. 806.

Thus, the positive findings in the short-term screening studies, which suggest a carcinogenic potential for Red No. 2, must be given considerable weight even though there are several negative tests. In view of the positive results, Red No. 2 should be regarded as a potential or actual mutagen, and thus a potential carcinogen. Mayer G–222, p. 38; Legator, G–223, pp. 20, 34.

E. *Conclusion on Carcinogenicity.* I find that the noncarcinogenicity of Red No. 2 has not been shown, due to the deficiencies in the Taylor/Monlux study and the Webb Rat Study, in combination with the questions of carcinogenicity arising from the Andrianova (Russian) study, and the lymphosarcoma finding in the Taylor/Monlux study singly and as reinforced by the total tumor findings in female rats in this test. These factors are sufficient to deny the petition for Red

No. 2. The results from the short-term screening studies provide additional support for that conclusion.

IV. Chemistry Issues

A. *Introduction.* In addition to providing a hearing on whether the data established that Red No. 2 is not a carcinogen, the Notice of Hearing provided a hearing on whether the statutory considerations support a determination that Red No. 2 is safe. 41 FR 29896, 29897 (July 20, 1976). At the hearing, the Bureau introduced evidence to show that Red No. 2 could not be considered safe due to questions about its purity, stability and metabolism. In its brief to the ALJ and in its exceptions to the Initial Decision, CCMA has maintained that the ALJ should have excluded the Bureau's evidence on these three matters. CCMA also argues that if the issues are reached, they should be resolved in CCMA's favor.

The issues concerning the purity, stability and metabolism of Red No. 2 have been characterized as the "chemistry issue," by CCMA. The Bureau regarded the purity issue as an aspect of the "carcinogenicity issue" rather than as part of the chemistry issue (Bureau Brief, pp. 48–50). All of the issues are discussed in this Decision as part of the chemistry issues because all of them are affected by similar considerations with respect to the appropriateness of a definite resolution in this proceeding.

Before discussing whether the chemistry issues should be resolved in this proceeding, I will summarize the Bureau's evidence with respect to each of these issues, and CCMA's reasons for believing the issues should be decided in its favor. Then, I will discuss the disposition made in the Initial Decision, CCMA's arguments for excluding the chemistry issues, the Bureau's reply, and my reasons for not resolving the issues at this point in this proceeding.

B. *Parties' Contentions on the Chemistry Issues.* 1. *Purity.* At the hearing, the Bureau introduced evidence about the presence of an impurity, alpha naphthylamine, in Red No. 2 and it primary intermediate, as detected by Canadian and FDA investigators. G–269, G–220. Although CCMA disputed the point (CCMA Brief, pp. 200–01), the Bureau introduced evidence showing alpha naphthylamine to be a suspect carcinogen. G–197, G–251–58. The Bureau stated that it was "unknown" whether the positive findings in the chronic feeding studies on Red No. 2 were due to the presence of this impurity. Bureau Brief, pp. 48–49.

CCMA suggested that Red No. 2 could be considered safe, even if alpha naphthylamine might be a carcinogen, because the manufacturers can "remove" the alpha naphthylamine down to any level set in FDA specifications, "even if the specifications were in parts per billion." CCMA Brief, p. 199. CCMA also noted that science cannot measure to zero to assure complete absence of a substance; "the best that science can do is to remove something to the extent that the level approaches zero." Id., P. 198 at n. 1 citing Tr. 652.

On the other hand, the Bureau maintained that "[t]he efforts made to eliminate alpha naphthylamine from Red No. 2 are interesting but, if they are of no avail, as they apparently have been, they are irrelevant in light of the Delaney Amendment." Bureau Brief, p. 12.

2. *Stability During Manufacturing and Storage.* The Bureau maintains that it is reasonable to require proof of Red No. 2's stability before ingestion under conditions of its most important uses in view of the showing that Red No. 2 is unstable in some foods under ordinary conditions of use, the sparse data on its stability in most uses, and the burden of proof on CCMA to show safety. Reply, pp. 49–50; Brief, pp. 93–96, 109.

The Bureau introduced evidence to show that Red No. 2 degraded substantially under ordinary conditions of use in chocolate pudding cookies, beef-flavored analogues, toffees and some other foods. G–5, G–7, G–10, G–17, G–221.

According to CCMA, it is "impossible to answer all questions regarding the stability of any color additive," but nonetheless a substantial amount of evidence showed that there were no legitimate concerns about this aspect of the safety of Red No. 2. Ex. p. 133. CCMA maintains that it is *undisputed* that in the color's "widest uses" (in soft drinks, frozen desserts, dairy products and candy), it is completely stable. Ex., p. 134 citing P–162, p. 8. The Bureau's stability evidence relates only to foods like baked cookies, beef analogues and pudding. Ex., p. 134. In its reply, the Bureau recognizes that Red No. 2 has been shown stable in "some" uses, but states that little information is available on "most" of its uses. Reply, p. 49. Until Red No. 2 is tested for stability "in all of its major proposed uses," it should not be approved generally. Id.

CCMA also argued that Red No. 2 would be safe, even if unstable, because breakdown occurs in the same way that it occurs during metabolism, (that is, at the azo linkage) and only harmless metabolites are produced. Ex., p. 135, citing G–221, p. 7 and P–157, p. 5). CCMA further argued that both halves of the azo linkage are sulfonated, and that, like all sulfonated compounds, the breakdown products are resistant to absorption in the bloodstream, are rapidly excreted from the body, and are rendered non-toxic. Ex., p. 135, citing P–107, p. 31; P–157, p. 37. The Bureau rebuts this arguments by pointing out that in the study cited by CCMA not all the breakdown products were identified. Reply, pp. 49–50; Brief, pp. 94–95, citing P–154, p. 14, P–157, p. 35.

CCMA also maintained that in the foods in which Red No. 2 breaks down it would break down in the same way in test animals and in humans. Therefore if ingestion of the color does not cause harm, as shown in tests on animals, "then the same breakdown which might occur in some few commercial uses would be equally harmless." Ex., p. 139.

According to the Bureau, however, even if animal tests showed Red No. 2 to be safe after ingestion, the tests would not be relevant to show that it is safe with respect to breakdown products prior to ingestion since the color may break down differently in the air than in the body. Brief, p. 95, citing P–157, p. 35; P–154, p. 14.

The Bureau also maintained that Red No. 2 had not been shown to be safe because studies showed that under some conditions Red No. 2 could "readily" desulfonate and could liberate alpha naphthylamine or other potentially harmful byproducts. G–1, G–221, pp. 11–14, P–157, pp. 30–31. The Bureau considered this evidence to be relevant to show concerns about both the stability and the metabolism of Red No. 2. Bureau Brief, pp. 95–96, Reply, p. 50. CCMA argued, though, that the Bureau's evidence to show desulfonation relies on a process that is unrelated to the actual conditions of use of the color and that did not in fact desulfonate the color. Ex., pp. 136–137, citing Tr. 599–657.

3. *Metabolism.* The Bureau maintains that the metabolism of Red No. 2. in animals and in humans is not understood. Reply, p. 50; Brief, pp. 96–97. In one study, researchers were able to recover only 69 percent of a dose administered orally to animals, thereby confirming that they do not fully know the metabolic pathways of Red No. 2 or its toxicity. Reply, p. 50; Brief, pp. 96–97, 109.

The Bureau urged the adoption of a finding that the metabolism of Red No. 2 provides no evidence of its safety because it is not well understood. Bureau Brief, p. 110. In its Brief, the Bureau maintains that Red No. 2 cannot be considered safe until the metabolites are "better identified" and the questions about its metabolism and the fate of

possible metabolites are satisfactorily resolved. Bureau Brief, p. 97. These questions apparently concern whether all the other metabolic by-products are excreted, as maintained by CCMA; whether Red No. 2 could be desulfonated by bacteria in the gut; and whether it could break down into several compounds, not all of which would be sulfonated. Reply, p. 50. CCMA regarded the possibility of desulfonation of Red No. 2 by bacteria in the gut as fairly "impossible." Ex., p. 138, citing Tr., p. 288.

C. *Initial Decision.* The ALJ noted that evidence concerning "the chemical properties" of Red No. 2 and "related chemicals" had been presented. Id., p. 18. He stated that the association of "related chemicals" to Red No. 2 "appears to occur" in three situations: (1) The dye "may contain some of its precursor chemical and breakdown products" before it is added to food; (2) the manufacturing processes and storage of food containing the color cause the formation of other chemicals from Red No. 2; and (3) Red No. 2 is modified or broken down when it is metabolized after ingestion. Id., p. 18.

The ALJ found that "consideration of the properties of Red No. 2 and its related chemicals raise additional unanswered questions as to the safety of the dye." Id., p. 19. The basis for this determination was the "suspicion" that certain metabolites of Red No. 2, formed by breakdown in the body after ingestion, are carcinogenic and that others are unidentified. Id., p. 18. The ALJ also referred to the "problems of degradation" of Red No. 2 during processing and storage, but it is unclear whether he regarded these problems as also creating "unanswered safety questions" or as ones that should be solved by the promulgation of definitive standards for the manufacture, storage, and use of Red No. 2. Id., pp. 18–19. The ALJ did not specifically rely on the purity issue.

D. *Exclusion of the Chemistry Issues; Exceptions.* CCMA objects to any consideration of the chemistry issues in this hearing because the issues were not specifically raised as a basis for the denial of the petition, and they were not set forth as issues for the hearing in the Notice of Hearing. Ex., pp. 43–58. Instead, the Bureau raised them at the Pre-Hearing Conference as issues to be considered as part of the general issue of the safety of Red No. 2 originally designated as a hearing issue at the request of CCMA. CCMA also maintains that adding the chemistry issues would violate the statutory procedures under which the Commissioner is to give his reasons for denying a petition before a hearing is held on objections to the decision. CCMA further argues that addition of the issues would also be inconsistent with the agency's regulations in 21 CFR 12.24 and 12.26 and the court decisions under which agency action must be justified on the grounds used by the agency and not on unasserted grounds.

In addition, CCMA maintains that it is improper to consider this issue in light of a statement in an FDA memorandum on the day after the hearing on Red No. 2 was granted by the then Commissioner. That memorandum stated that:

> Questions concerning stability cannot be answered with the current state of the act (sic) and, therefore, should not be a reason for not listing a color additive. It was agreed that listing regulations would be prepared for those color additives where stability under the conditions of use was the only remaining question to be answered. Ex., Attachment A.

CCMA believes that it would be "an unconscionable denial of fundamental fairness" to require it to resolve questions that cannot be resolved with the current state of the art. Ex., p. 58.

E. *Inclusion of Chemistry Issues; Bureau Reply.* The Bureau sets forth numerous reasons why the chemistry issues were proper issues in the hearing. Among other arguments, the Bureau maintains that if chemistry was an improper matter for consideration at the hearing CCMA should have raised an objection at the prehearing conference. CCMA's failure to object until its Brief before the ALJ was "much too late." Reply, p. 43. CCMA itself raised the chemistry issues in its objections to the Commissioner's order when it argued that the Commissioner failed to consider the statutory requirements of safety set forth in 21 U.S.C. 376(b)(3)(A), and that section requires the stability be considered an issue. Id., p. 44. The Bureau also argues that the stability issue is directly related to the cancer issue, which CCMA agrees is a proper issue in the hearing. Id., p. 45.

The FDA document cited by CCMA to show that stability was never intended to be an issue is double hearsay and is unclear as to whether it was intented to suggest that "stability could not be considered along with other safety considerations at a hearing even though not serving as a separate rationale for not listing a color." Id., p. 48.

Section 371(c)(2) of 21 U.S.C., cited by CCMA, does not state that the hearing may consider only those matters on which objections are filed. Id. Finally, the Bureau urges that notice is the only relevant issue, and that CCMA had ample notice and in fact introduced witnesses who testified on the stability issue. Id.

F. *Analysis.* I believe there is no legal obstacle to consideration of the chemistry issues, in all their aspects, even though they were not included as issues in the notice of hearing. However, even though it would be permissible to consider them, I believe I should not consider them here because it is unnecessary to do so, and because the issues need further examination on a better record and in a broader public proceeding.

I am not persuaded by the arguments of CCMA that the agency is precluded from considering a new issue at a formal hearing solely because it was not specifically included in the notice of hearing. The issues for the hearing were broadly stated, and the specific issues were expressly added by the presiding officer. FDA's current procedural regulations (21 CFR 12.35(b)) permit the presiding officer to add issues to those in the notice of hearing. Moreover, CCMA apparently agreed to the addition of the chemistry issues at the Pre-Hearing Conference, (Tr., p. 59, August 9, 1976) and fully litigated the issues. Consequently, the introduction of the issues into the hearing involved no lack of notice or other procedural unfairness to CCMA.

The Administrative Law Judge correctly rejected CCMA's argument that the hearing must review only the ground expressly asserted by the agency in the Notice of Hearing. Id., p. 6. The agency may rely on grounds not in the original complaint when the party has had notice and a full opportunity to defend. *Golden Grain Macaroni Co.* v. *FTC.* 472 F.2d 882, (5th Cir., 1972), cert. denied, 412 U.S. 918 (1973).

The court cases cited by CCMA, such as *SEC* v. *Chenery Corp.*, 318 U.S. 80 (1943), are not analogous. The cases require an agency to defend agency actions upon judicial review on the grounds articulated by the agency, rather than on post hoc rationalizations first stated in court after the completion of the administrative process. CCMA would apply that principle to preclude an agency from relying on any basis for an administrative decision that had not been expressly stated in the initial notice of hearing, even though the parties had been informed of the specific issue before the hearing and had litigated its relevance to the decision in the administrative process. The proposed extension of the *Chenery* line of cases is unwarranted. The purpose of the *Chenery* rule is to provide the court with the agency's contemporaneous reasons for its action. That purpose has no application to the

administrative process within an agency.

CCMA also cites in support of its position judicial decisions that an agency may not "change theories in midstream without giving respondents reasonable notice of the change." Ex., p. 56, citing *Rodale Press Inc.* v. *FTC.* 407 F.2d 1252, 1256 (D.C. Cir. 1968). Leaving aside the question of whether the agency has changed theories, it is clear that CCMA has had reasonable notice of the chemistry issues, and therefore these cases do not support its position.

The statute should not be interpreted as barring the addition of hearing issues. To do so would preclude the addition of issues that might be important to a sound decision, and that might eliminate the need for a subsequent, largely repetitive hearing. The added issue may grow out of the issues already present. The CCMA argument that FDA's regulations in 21 CFR 12.24 and 12.26 preclude addition without prior consideration by the Commissioner of the objections is in error in view of the clear provision in 21 CFR 12.35(b) permitting the addition of issues, and the fact that 21 CFR 12.26 relates to modification of a rule to eliminate or alter the need for a hearing, and not to the addition of issues after a hearing begins.

Nonetheless, although additional hearing issues "may" be added by the presiding officer, it is not always reasonable and appropriate to add them. I believe that it is not necessary to consider the chemistry issues at this point because the absence of chronic feeding studies adequate to show the safety of Red No. 2 provides, as already discussed, a sufficient basis for denying the petition for permanent listing of Red No. 2 as a color additive. Consequently, it is unnecessary to decide whether the chemistry issues provide separate grounds for denying the petition. If further testing is done that shows Red No. 2 to be safe, it will then be necessary to consider whether there are additional grounds, on the basis of chemistry or any other factor, for denying the petition.

The chemistry issues involve complicated and important matters, on which agency policy is still developing. These questions include the specifications to be set to detect impurities in a color additive when an impurity is a suspect carcinogen and a starting material in making the color, and the overall treatment of nonfunctional constituents of food and color additives. Because these matters were not set as hearing issues, the Commissioner did not need to state before the hearing the agency's position on them. Once the notice of hearing was issued, the agency's regulations imposed a separation of functions requirement, which precluded discussion of the policy issues between the Bureau and the Commissioner. 21 CFR 10.55. As a result, this proceeding did not have the benefit of the usual discussion that occurs when the agency is developing its policy in a new area prior to a hearing.

The record in this proceeding is limited to Red No. 2, and consequently the impact of any general resoluton of the chemistry issues upon other color additives is not apparent from this record. I am reluctant to establish precedent that could have a broad impact on many matters outside this record when it is not necessary to do so, and when deferral of this issue will permit both broader public comment and further deliberation within the agency on the general policy issues.

Moreover, these issues, important as they are, appear to have played a minor role in the hearing. The record is sketchy and less focused and informative than is desirable for a definitive resolution of these matters. For example, it is not clear whether the Bureau believes full metabolic data would be necessary if there were an adequate chronic feeding study that showed the safety of Red No. 2. The Bureau urged the ALJ to adopt a finding that the metabolism of Red No. 2 is not well understood, and, therefore, "provides no evidence of safety." Bureau Brief, p. 110. Although a lack of understanding of metabolism cannot contribute to a showing of safety, this proposed finding seems to stop short of maintaining that gaps in metabolic data necessarily and independently preclude a finding that a substance is safe.

Analogously, on the impurity issue, CCMA has suggested, as stated above, that Red No. 2 could be considered safe even if it is made from a starting material that may be a carcinogen because the manufacturers can "remove" the *alpha*-naphthylamine down to any level set in FDA specifications, "even if specifications were in the parts per billion." CCMA brief, p. 199. In its Brief to the ALJ, the Bureau stated that the efforts to eliminate *alpha*-naphthylamine are "interesting, but, if they are of no avail . . . they are irrelevant in light of the Delaney Amendment." Bureau Brief, p. 12.

Although the Bureau's position may be the correct one for the agency to adopt, because this issue of the level of specifications for impurities can arise with other colors, it is preferable to resolve this issue in a separate proceeding that permits broader public scrutiny.

Similarly, the stability issue is not ripe for decision. The Bureau has shown that certain uses of Red No. 2 are unstable, and states that other uses are stable but does not identify them. CCMA maintains that it is "*undisputed*" that Red No. 2 is stable in its "widest uses." Ex., p. 134. It is unclear whether or not the Bureau agrees with this statement. The issue is too important to resolve definitely on a record that is so unclear. The Bureau called for stability testing of Red No. 2 "in all of its most important uses" Reply, p. 49. Although some further testing appears desirable, I am unsure which food uses would be covered by the finding that more testing is needed and in which uses Red No. 2 has already been shown to be stable. Furthermore, CCMA suggests that changes in manufacturing practices and the conditions of use could eliminate the stability problems. Ex., p. 134, n. 70.

The Bureau also introduced evidence to show that Red No. 2 can break down before ingestion in a way that desulfonates it and may produce harmful by-products. G–1. CCMA maintained that the process of desulfonation used in this test is not reasonably related to the conditions in which Red No. 2 is used in food. Ex., p. 136, citing Bureau witness Link, Tr., pp. 606–57. The Bureau's response to this contention is not clear from the record. A Bureau witness also suggested that color changes in certain foods indicated desulfonation could be occurring (Link, G–221, pp. 13–14), but CCMA suggested another explanation for the color change. CCMA Brief, p. 193 citing G–22, Link, Tr., p. 623. I do not see in the record an explanation of why this particular color change is associated with desulfonation. It may be that the Bureau could make a better showing on this matter if it became the focused basis for a possible denial of approval. Consequently, I will refrain from making any further conclusions on this issue because the record is inadequate to resolve it with confidence.

For the reasons stated, I do not decide the chemistry issues, including the issues of stability, metabolism and purity.

V. *Topical Uses of Red No. 2 in Cosmetics.* A. *Initial Decision and Record.* The Commissioner's order denying the petition for permanent listing of Red No. 2 applied to its "use in food, drugs, and cosmetics." 41 FR 15053 (April 9, 1976). The original petition (CAP No. 36) filed in 1968 sought approval for use in foods, drugs, and cosmetics. Docket 76C 0033, No. 4, Letters to the Commissioner from Hazleton Laboratories, Inc., September

16, 1965. That the topical use of Red No. 2 was intended to be included in that petition is evidenced by a "Proposed regulation" attached to the petition and made part of it by reference wherein it is stated:

(b) Uses and Restrictions: The color additive is safe for general use in coloring foods, ingested or topically applied drugs, lipsticks, and other cosmetics * * *.

In addition, in a May 14, 1974 letter to the Commissioner, the CTFA, successor organization to one of the petitioners, amended the proposed regulations contained in the petition to state:

The color additive is safe for general use in coloring topically applied cosmetics including those intended to be used in the area of the eye.

Docket 76C 0033, No. 4, letter to the Commissioner from Cosmetic Toiletry, and Fragrance Assn., May 14, 1974. It is thus clear that the petition (CAP No. 36) included topical uses of Red No. 2 and that when the Commissioner denied the petition for the permanent listing of Red No. 2 as a color additive for use in foods, drugs, and cosmetics, his order included topical uses of the color.

CTFA objected to the denial of the petition to the extent it related to use in cosmetics applied topically, but CTFA did not request a hearing because it intended to file a new petition for use of Red No. 2 in externally applied cosmetics and request a hearing if that petition were denied. The Commissioner stated, however, that the CCMA's objections and request for a hearing concerning the Commissioner's finding that Red No. 2 has not been shown to be safe raised the color's safety for all intended uses. The Commissioner concluded that the issues raised by CTFA's objections would necessarily be aired at the evidentiary hearing and that a "separate and repetitive hearing on the same issues" would not be granted if a new petition were filed. 41 FR 29896, 29897 (July 20, 1976). The Commissioner granted a hearing on, among other things, whether Red No. 2, if not approvable for all petitioned uses, may be approvable for certain limited uses. Id.

Both CCMA and CTFA appeared at a prehearing conference and objected to the inclusion of topical uses as an issue. Transcript of Pre-Hearing Conference at 6–8, August 9, 1976. In a pre-hearing order, the ALJ ruled that the safety of the topical use of Red No. 2 was properly an issue in the hearing. Id. at 9–10.

Neither CCMA nor CTFA took an interlocutory appeal from this order, as each could have done. See 21 CFR 2.89 (1976) as modified by order of the ALJ. Id. at 5. CTFA later withdrew from the proceeding and stated that it was unnecessary for it to appear with respect to ingested cosmetics and that consideration in the evidentiary hearing of topical uses would be "inappropriate and unlawful." In its Withdrawal of Appearance, CTFA stated that CCMA's objections related to the other uses of Red No. 2 had been withdrawn to the extent they could be construed as applying to topical uses. Furthermore, CTFA argued that the safety of Red No. 2 for topical uses had not been properly evaluated by FDA, and the FDA denial of listing "was based solely on safety questions allegedly raised by studies relating to ingestion of the additive." Docket 76C 0033, No. 29, Withdrawal of Appearance of CTFA, p. 2.

The Initial Decision does not discuss the applicability of the proceeding to topical uses, but finds that Red No. 2 cannot be approved for certain limited uses (in effect prohibiting topical uses), and recommends an order that would completely deny the petition for permanent listing. Id, p. 20–21. No exceptions were taken on this point, but the Bureau in its reply suggested that it would be appropriate to discuss the evidence on topical uses in this decision. Reply, p. 3 n. 1.

The Bureau was the only participant to present information at the hearing about topical uses. Testimony indicated that Red No. 2 used in cosmetics would be expected to go through the skin but the amount could not be estimated without an experiment. Scheuplein, G–225, p. 10–11. Other evidence related to the inadequacies of animal skin painting studies on Red No. 2 to show its safety with respect to the risk of skin cancer. Weinberger, G–218, p. 20. CCMA introduced no evidence on topical uses and even proposed a conclusion of law that "no evidence was introduced in this hearing on the topical uses of FD&C Red No. 2, and accordingly, it was not shown to be safe for this use." CCMA Brief, p. 226.

B. *Analysis.* As indicated earlier, the order denying the listing of Red No. 2 applies to all its uses, including topical uses. 41 FR 15053 (April 9, 1976). Consequently, the order issued by the Commissioner applies to the topical uses of Red No. 2, and interested persons had an opportunity to obtain a hearing on the appropriateness of that order. Since no evidence was introduced at the hearing to show that Red No. 2 should be permitted to be used in topical applications, there is no reason to alter the original order that denies completely the petition to list permanently Red No. 2. In this proceeding, I have determined that Red No. 2 cannot be approved for use as a color additive because there is no adequate study that shows with reasonable certainty that it is not a carcinogen. I believe that a repetitive hearing on the same issues should be avoided in the future.

VI. Findings of Fact

Cancer

1. The Webb rat study (P–2) is deficient and cannot be considered as important evidence of the safety of Red No. 2 for the following reasons:

a. Insufficient histopathologic examination was conducted or animal tissues in this study. G–11, p. 51; Borzelleca, T. 114; Weinberger, G–0272, G–218, p. 14.

b. The gross pathology in the study was not acceptable because it was performed on shruken, hardened, and discolored tissues, which had been fixed in formalin. Weinberger, G–218, p. 14.

c. The test animals suffered from intercurrent disease, which was unrelated to the test chemical. This fact made attribution of any effect or lack thereof to Red No. 2 impossible. Smuckler, G–217, p. 15.

d. Due to low survivorship and a high rate of autolysis, the number of animal tissues examined in the study was insufficient to provide confidence that the study would have detected a carcinogenic effect of the chemical. Weinberger, G–218, p. 15; Smuckler, G–217, p. 15.

2. The 1961 Webb mouse study (P–3) cannot be deemed proof of safety of Red No. 2 for the following reasons:

a. Low survivorship and autolysis lowered the number of observable tissues to the extent that the study was unable to detect any but the stronger carcinogen. Weinberger, G–218, pp. 17–18; Schneiderman, G–227, p. 11.

b. The quality of the gross pathology in this study was unacceptable. Weinberger, G–218, p. 17.

c. The histopathology performed following sacrifice was inadequate. Weinberger, G–218, p. 17; Smuckler, G–217, p. 19.

d. Intercurrent disease, including parasites, pervaded the mouse population and thereby rendered impossible attribution of any effect or lack thereof to the chemical. Smuckler, G–217, p. 20.

3. The dog study (P–5) cannot be given any weight in proving the safety of Red No. 2 due to the small number of dogs administered the compound in the study. Gaylor, G–226, p. 30; Weinberger, G–218, p. 18; Smuckler, G–217, p. 8. Furthermore, the dog is not a good animal to use in such a study unless the

purpose of the study is specifically to look for bladder tumors. Smuckler, G–217, p. 9.

4. The results from the FDA Taylor/Monlux chronic feeding study (P–6, P–7, 1975) provide strong presumptive evidence of the carcinogenicity of FD&C Red No. 2. Gaylor, G–226, pp. 18 ff; Schneiderman, G–227, pp. 17–18.

5. In the Taylor/Monlux chronic feeding study there was a dose response concerning lymphosarcomas. Gaylor, G–226, pp. 25–6; Schneiderman, G–227, p. 18.

6. The so-called mix-up in the Taylor/Monlux chronic feeding study could only have meant that the true cancer rate caused by Red No. 2 may be greater than, but could under no circumstances be less than, that shown in the study. Gaylor, G–226, pp. 26, 41–43. The extent and impact of the mix-up cannot be determined with reasonable confidence. P–41, Minutes of Toxicology Advisory Comm., p. 4 (March 8–9, 1976).

7. The deficiencies in the Taylor/Monlux chronic feeding study, to the extent that they effectively decreased the number of tissues that could have been examined histopathologically after 24 months, could only cause greater concern about the carcinogenicity of Red No. 2, because the smaller a study that picks up a positive effect, the more potent that effect must be. Gaylor, G–226, p. 39; Schneiderman, G–227, p. 13.

8. The variety of tumors seen in the Taylor/Monlux chronic feeding study raise a suspicion of cancer and preclude a finding based on that study that the substance is not a carcinogen. Farber, G–219, p. 7.

9. The Taylor/Monlux chronic feeding study is inadequate to evaluate the safety of Red No. 2 due to the high incidence of intercurrent renal disease. Smuckler, G–217, p. 23.

10. The finding of a dose response in lymphosarcomas in the FDA chronic feeding study is important in light of the finding of lymphosarcomas in the Andrianova study. (G–14).

11. Simultaneous statistical inference is an inappropriate technique to use to analyze dose-response data because it ignores relevant biological information (Gaylor, G–226, p. 27, Tr., p. 949), and may increase the risk of a false negative to unacceptable levels (Tr., p. 971).

12. A one-tail test is more appropriate than a two-tail test for analyzing data from experiments to determine safety because a one-tail test will give the maximum information about the safety of the test compound. Schneiderman, G–227, p. 19.

13. The Russian study by Andrianova (G–14) supports the conclusion that FD&C Red No. 2 has not been shown with reasonable certainty not to be a carcinogen because the Russians obtained a highly statistically significant positive result for the amaranth, tested against both positive and negative controls. Gaylor, G–226, p. 29.

14. It is uncertain whether the Andrianova study tested a substance that is chemically equivalent to FD&C Red No. 2. P–16, pp. 13–14, P–145.

15. The other studies by Mannell and Grice (P–1), and Nelson (G–270) do not provide evidence of the safety of Red No. 2 because, inter alia, the studies were too small to allow confidence in a negative result. P–150(2), p. 52, G–208, p. 9.

16. One cannot conclude from the fact of long-time use of FD&C Red No. 2 in the food supply that it is safe. Legator, G–223, pp. 30–31.

Mutagenicity

17. A chemical that is a mutagen is also a potential carcinogen. G–222, p. 35; G–223, pp. 28–29; G–224, pp. 4–5; P–151, pp. 16, 30–31, 35, Tr., pp. 867–68.

18. Results achieved in cytogenetics experiments, a short-term screening study for mutagenic potential, indicate that Red No. 2 and a known metabolite are potential carcinogens.

a. Dr. Legator, witness for the Bureau, is a recognized expert in in vivo cytogenetic testing. G–223, pp. 6–7.

b. The in vivo cytogenetics experiment performed by Frances Moreland produced positive cytogenetic effects for the R-amino salt metabolite of Red No. 2 and for Red No. 2 itself. G–3, p. 4; G–223, p. 17; Tr., p. 792.

c. The in vitro anaphase cytogenetics test performed by the Stanford Research Institute also showed a cytogenetic effect caused by Red No. 2. P–90, pp. 36, 38; G–223, p. 19; Tr., p. 801.

d. The use of intraperitoneal administration in an in vivo cytogenetics experiment is appropriate (Tr., p. 806; G–223, pp. 15, 16), and in fact is advisable because it maximizes the chance of observing a positive effect. Tr., pp. 805–806.

e. Chromatid and chromosomal breaks, observed in the cytogenetic studies, are clear indications that the substances tested cause cytogenetic effects, since any substance that causes chromatid breaks will also cause the other types of aberrations observable in cytogenetic experimentation. Tr., pp. 786–787, 798, 857.

f. Dr. Brusick's criticism of the in vivo cytogenetics test results on the ground that the concurrent controls used for evaluation had an abnormally low spontaneous rate of cytogenetic effects, as compared to an average spontaneous rate in different strains in historical controls, as observed in other laboratories (Tr., pp. 430–31), is invalid. There is a great deal of variability between strains (Tr., pp. 428, 870) and between laboratories (Tr., p. 439) and between particular effects (Tr., p. 437) in evaluating the results of cytogenetic tests (Tr., p. 443). The historical control results against which Dr. Brusick was measuring concurrent controls in several tests were from averages or ranges using unspecified strains or sublines observed in different laboratories under unspecified conditions (P–151, pp. 42–43) or reflected a range of historical spontaneous background aberrations observed in other laboratories under unspecified conditions (P–151, p. 48). These are invalid comparisons for the purpose of establishing and utilizing a control for the evaluation of data from a particular experiment (Tr., pp. 77, 439). He also agreed that if a particular strain had a consistently low control level, the use of that control level value is proper. (Tr., p. 870).

g. The positive results achieved in the cytogenetics experiments suggest that a test substance, in this case, FD&C Red No. 2, is in fact a carcinogen. Tr., pp. 862, 868, 790–91, 795–96.

19. Mutagenesis testing in drosophila produced strong evidence that Red No. 2 is a mutagen.

a. The two drosophila tests that did not produce positive results with Red No. 2 do not invalidate the strong positive achieved in the Oster study because more than two non-positive tests would be required to negate a positive test even if the conditions under which all were performed were identical (G–227, p. 24; Tr., p. 669); and because (1) the test strains used in the three studies were different (G–222, p. 28) and (2) the test flies used in the Oster study were starved, providing a certainty, not present in the other two tests, that the flies actually consumed the Red No. 2 used (G.–222, p. 28).

b. Red No. 2 produced clearly positive mutagenetic results in two tests run by Dr. Oster (G–4, p. 1; G–226, pp. 44–45; G–222, pp. 21, 26–27), of particular importance in light of the admitted expertise of Dr. Oster in drosophila testing. P–151, p. 43.

c. The charge that the Oster study was too small is invalid in light of the fact that the statistical analysis of the study took into account the study size and still found a significant positive result. Tr., pp. 415, 443.

d. The Oster study results were based upon a valid concurrent control, which was in line with other control figures achieved by Dr. Oster in the same laboratory at about the same time, the most appropriate control group to use

for purposes of comparison (Tr., p. 442) in light of the fact that strains may vary in the levels of spontaneous mutations they produce (Tr., pp. 870, 872). The finding is particularly significant in light of the fact that statistical significance was found when the positive results observed were compared with the combined negative control figures for all of the experiments that Dr. Oster performed. Tr., p. 675.

e. In light of the fact that the sex-linked recessive lethal test is more sensitive than the other tests run by Dr. Oster (Tr., pp. 434, 667–68; G–185, pp. 128–29), it is not significant that the less sensitive tests performed by Dr. Oster do not show a mutagenetic effect for Red No. 2. Tr., pp. 667–68.

20. The R–amino salt of Red No. 2 produced a clearly positive result in the specific locus assay utilizing L5178Y mouse lymphoma.

a. The results achieved in the mouse lymphoma assay were clearly positive (G–165; G–223, p. 23; G–151, p. 70). If the mutagenetic effect observed is caused by the quinone of the R-amino salt, cause for concern remains since quinones are known to be formed in the body. G–165.

b. Results in the mouse lymphoma assay appear to correlate well with carcinogenesis. P–151, pp. 73–74.

21. The positive result observed in the dominant lethal assay, particularly in light of the insensitivity of that test, shows that Red No. 2 is mutagenic.

a. The dominant lethal assay is considered to be the least sensitive mutagenicity test. G–223, p. 22, Tr., p. 468.

b. A positive effect caused by the R-amino salt of Red No. 2 was observed in a dominant lethal assay. G–165; G–226, pp. 51–52.

c. Dr. Brusick's criticism of the positive finding as being caused by an abnormally low rate of spontaneous dominant lethal effects in the concurrent control is unsupported and may be the result of comparison with historical controls, which can vary in incidence of spontaneous dominant lethal effects due to several factors, including the solvent used. Tr., pp. 436, 477–78.

d. Dr. Brusick's criticism of the positive dominant lethal study as using too toxic a dose in the highest level of administration is speculative. Tr., p. 472.

e. Dr. Brusick's suggestion that the dominant lethal assay should be used as a "risk evaluation test" and his contention that the dominant lethal assay should use the relatively insensitive ingestion route of administration are not accepted in light of the fact that the dominant lethal assay is relatively insensitive (Tr., p. 468) and that, by Dr. Brusick's own admission, the finding of a negative in a dominant lethal test provides no assurance that the chemical tested is not a carcinogen. Tr., p. 868.

22. The failure of Red No. 2 and four of its known metabolites to produce a positive effect in the Ames bacteria test does not indicate that Red No. 2 is not a mutagen.

a. A negative response in the Ames test should not be taken as an indication that the chemical involved is clearly not a mutagen. G–224, p. 6; G–223, pp. 20, 25–26.

b. Since Red No. 2 is metabolized in the gut rather than in the liver (G–224, p. 8; P–151, p. 62), Red No. 2 may well need to be metabolized before it causes a carcinogenic effect (P–151, pp. 34–35), and would not be expected to cause a mutagenic effect in the Ames test, even if it were a mutagen and carcinogen. P–151, p. 35; G–224, p. 8.

c. The testing of the known metabolites of Red No. 2 in the Ames test does not show that Red No. 2, when metabolized in the body, would not be mutagenic because it is not known what all of the metabolites of Red No. 2 are. Tr., pp. 410, 505; G–224, p. 8; G–223, p. 25.

d. Because Red No. 2 requires bioactivation in an organ other than the liver, the Ames test is insensitive test for determining its mutagenicity. Tr., 412; G–223, p. 24–25; P–159, p. 6.

23. The yeast studies done with Red No. 2 and its metabolites do not show them not to be mutagens.

a. The yeast study may have the same need for proper bioactivation as the Ames test. G–222, p. 32.

b. Yeast cells are not as permeable to larger molecules as are, for instance, the modified bacteria used in the Ames system and thus may be less likely to show a mutagenic effect. G–222, p. 32; Tr., p. 500.

24. No other study shows Red No. 2 not to be mutagenetic.

a. Neither the test for teratogenicity nor the multi-generation reproductivity study is an acceptable study for mutagenicity, since it would be very unlikely that anyone would observe a mutation in them. Tr., p. 427.

b. The host-mediated assay is so insensitive that it has been discontinued as a screening tool in many testing laboratories. P–151, p. 26.

25. CCMA's suggestion that the positive results achieved with the R-amino salt in the Moreland cytogenetic tests, the FDA dominant lethal test, and the mouse lymphoma test could be ascribed to an impurity in the sample used is speculative (Tr., pp. 856–57) and conflicts with what evidence there is on this subject. Tr., pp. 808–09.

26. The range of results achieved with Red No. 2 in mutagenicity testing is not unusual for an active mutagen (G–223, p. 20; Tr., pp. 434–35; G–224, p. 7), and it is thus necessary to consider the entire set of results achieved. G–224, p. 6; G–223, pp. 26–27.

27. Expert opinion indicates that Red No. 2 has been shown by mutagenicity testing to be a potential carcinogen. Mayer, G–222, p. 38, Legator, G–223, p. 34.

Topical Uses

28. Petitioner submitted no evidence of safety of Red No. 2 for topical uses.

29. Red No. 2, when applied topically, is reasonably expected to be absorbed through the skin, although the amount that would be absorbed cannot be estimated without an experiment. Scheuplein, G–225, pp. 9–11.

30. The skin painting studies (G–118, 119) cannot serve as evidence of Red No. 2's safety in regard to skin cancer, since their duration was short (90 days), survivorship of the animals was low, and autolysis was unacceptably high. Weinberger, G–218, p. 20.

VII. Conclusion of Law

1. The burden of proof in this hearing rests upon the Petitioner, who must prove safety on all issues to a reasonable certainty.

2. Petitioner has not proven to a reasonable certainty that FD&C Red No. 2 is not a carcinogen in man or animals.

3. Petitioner has not proven to a reasonable certainty that FD&C Red No. 2 is safe for topical uses.

VIII. Order

The petition for permanent listing of FD&C Red No. 2 (CAP 36) is denied.

In accordance with sections 701(e)(3) and 706(d) of the act (21 U.S.C. 371(e)(3) and 376(d)) this order is made effective April 24, 1980.

Dated: January 15, 1980.

Jere E. Goyan,

Commissioner of Food and Drugs.

[FR Doc. 80–1911 Filed 1–23–80; 9:35 am]

BILLING CODE 4110–03–M